JPL Publication 91-44

Communications and Control for Electric Power Systems

The 1990 Report

H. Kirkham
A. Göttsche
D. Niebur
H. Friend
A. Johnston

January 15, 1991

Prepared for

Office of Energy Management Systems
United States Department of Energy
Through an agreement with

National Aeronautics and
Space Administration
by

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No. 91-44	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Communications and Control for Electric Power Systems The 1990 Report		5. Report Date January 15, 1991 6. Performing Organization Code
7. Author(s) H. Kirkham, A. Gottsche, D. Niebur, H. Friend, A. Johnston		8. Performing Organization Report No.
9. Performing Organization Name and Address JET PROPULSION LABORATORY California Institute of Technology 4800 Oak Grove Drive Pasadena, California 91109		10. Work Unit No. 11. Contract or Grant No. NAS7-918 13. Type of Report and Period Covered JPL Publication
12. Sponsoring Agency Name and Address NATIONAL AERONAUTICS AND SPACE ADMINISTRATION Washington, D.C. 20546		14. Sponsoring Agency Code RE152 0203-PL-778-12-01-02-03
15. Supplementary Notes		

16. Abstract The first section of the report describes the AbNET system, a hardware and software communications system designed for distribution automation (it can also find application in substation monitoring and control.) The topology of the power system fixes the topology of the communications network, which can therefore be expected to include a larger number of branch points, tap points and interconnections. These features make this communications network unlike any other. The network operating software has to solve the problem of communicating to all thenodes of a very complex network, in as reliable a way as possible even if the network is damaged, and it has to do so with minimum transmission delays and at minimum cost. The design of the operating protocols is described within the framework of the seven-layer Open System Interconnectio hierarchy of the International Standards Organization.

Section 2 of the report describes the development and testing of a high voltage sensor based on an electro-optic polymer. The theory of operation is reviewed. Bulk fabrication of the polymer is discussed, as well as results of testing of the electro-optic coefficient of the material. Fabrication of a complete prototype sensor suitable for use in the range 1-20 kV is described. The electro-optic polymer is shown to be an important material for fiber optic sensing applications. Appendix A is theoretical support for this work.

The third section of the report presents the application of an artificial neural network, Kohonen's self-organizing feature map, for the classification of power system states. This classifier maps vectors of an N-dimensional space to a 2-dimensional neural net in a nonlinear way preserving the topological order of the input vectors. These mappings are studied using a nonlinear power system model.

17. Key Words (Selected by Author(s))	18. Distribution Statement Unclassified; unlimited		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 104	22. Price

JPL Publication 91-44

Communications and Control for Electric Power Systems

The 1990 Report

H. Kirkham
A. Göttsche
D. Niebur
H. Friend
A. Johnston

January 15, 1991

Prepared for

Office of Energy Management Systems
United States Department of Energy
Through an agreement with

National Aeronautics and
Space Administration
by

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Prepared by the Jet Propulsion Laboratory, California Institute of Technology, for the U.S. Department of Energy through an agreement with the National Aeronautics and Space Administration.

This publication reports on work performed under NASA Task RE-152, Amendment 203 and sponsored through DOE/NASA Interagency Agreement No. DE-AI01-79ET 29372 (Mod. A009).

ABSTRACT

The first section of the report describes the AbNET system, a hardware and software communications system designed for distribution automation (it can also find application in substation monitoring and control.) The topology of the power system fixes the topology of the communications network, which can therefore be expected to include a large number of branch points, tap points and interconnections. These features make this communications network unlike any other. The network operating software has to solve the problem of communicating to all the nodes of a very complex network, in as reliable a way as possible even if the network is damaged, and it has to do so with minimum transmission delays and at minimum cost. The design of the operating protocols is described within the framework of the seven-layer Open System Interconnection hierarchy of the International Standards Organization.

Section 2 of the report describes the development and testing of a high voltage sensor based on an electro-optic polymer. The theory of operation is reviewed. Bulk fabrication of the polymer is discussed, as well as results of testing of the electro-optic coefficient of the material. Fabrication of a complete prototype sensor suitable for use in the range 1-20 kV is described. The electro-optic polymer is shown to be an important material for fiber optic sensing applications. Appendix A is theoretical support for this work.

The operating point of a power system can be defined as a vector whose components are active and reactive power measurements. If the security criterion is prevention of line overloads, the boundaries of the secure domain of the state space are given by the maximum permissible currents of the transmission lines. The third section of the report presents the application of an artificial neural network, Kohonen's self-organizing feature map, for the classification of power system states. This classifier maps vectors of an N-dimensional space to a 2-dimensional neural net in a nonlinear way preserving the topological order of the input vectors. Therefore, secure operating points, that is vectors inside the boundaries of the secure domain are mapped to a different region of the neural map than insecure operating points. These mappings are studied using a non-linear power system model. Choice of security criteria and state space are discussed.

ACKNOWLEDGMENTS

The authors would like to thank Shannon Jackson for his continued support and assistance in the laboratory. It seems we have been acknowledging Shannon's help for a number of years—help in designing and making field sensors, fiber optic systems, and electronics. During 1990, Shannon became something of a chemist, as the electro-optic polymer voltmeter was developed. Joe Perry and Kelly Perry also deserve our thanks for their help with the chemistry of this device. Annie Aroyan deserves our thanks for her help as a consultant on the production of the manuscript.

Outside JPL, the two faculty advisors of the two graduate student authors deserve special mention. First, Prof. Ole Tønnesen of the Technical University of Denmark is thanked for helping arrange for Allan Göttsche's visit to JPL, and for his help and support of the work done here. The help and guidance of Prof. Alain Germond of the Ecole Polytechnique Fédérale de Lausanne in Switzerland is also acknowledged. He was in contact throughout the neural network project described here.

ABOUT THIS DOCUMENT

This report contains rather more mathematics than has been usual in the output of the Communications and Control Project. The job of typing it was started on a PC-type computer using an early version of WordPerfect, with the output to an early laser printer. To accommodate the mathematics, the text manuscript preparation program TEX was used. TEX is a trademark of the American Mathematical Society.

PREFACE

This report is the latest in a series of documents produced by the staff of the Communications and Control Project at JPL. External publication of this report, like that of its predecessor, was delayed by the need to file patents on work described in it. Its authorship differs from that of its predecessors in that two of the writers were graduate students pursuing advanced degrees at foreign universities. When this report was written, Allan Göttsche was a Ph.D. student at the Technical University of Denmark in Lyngby, near Copenhagen. His advisor was Prof. Ole Tønnesen. Dagmar Niebur was a German citizen working towards a Ph.D. at the Swiss Federal Institute of Technology in Lausanne, Switzerland. Her advisor there was Prof. Alain Germond.

The subject of Allan's work was decided upon jointly by Allan, Ole, Alan Johnston and me. The development of an optical voltmeter based on an electro-optic polymer seemed like an interesting topic, and one that could be accomplished during Allan's 10-month stay at JPL. The theoretical and practical work described in the second section of this report was later included in Allan's thesis.

The idea of using a Kohonen network to classify the security states of a power system had been selected as Dagmar's thesis topic before she came to the United States. To the Communications and Control Project (though not to JPL), this was a new field of study. The work described in the third section of this report provides a tutorial introduction to the topic. It is likely that, after Dagmar's return to Switzerland in 1992 we will continue to work in the area of artificial neural networks.

The work done in 1990 on AbNET, the fiber optics communications system designed at JPL for the distribution system, is described in the first section of the report. In order to speed the development in the forthcoming year, it was decided to base the hardware for an AbNET demonstration on Ethernet. The reasons for that decision, and its impact on our work are discussed.

The 1990 report, then, covers three topics: *communications* for the distribution system and for substations, *measurement* of the kind that might provide data at low cost for distribution automation, and *control* of the sort that occurs in the Energy Management Center. The common thread that links all the parts is, of course, the power system. In recent years, we have concentrated on communications and measurements. In effect, our return to the control aspects represents a return to power systems, for while communications systems and measurement methods can be developed in isolation, control of this very large, non-linear plant is the very heart of power systems work. We continue to be optimistic about the future challenges.

HAROLD KIRKHAM
PASADENA, 1991

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SECTION 1.

AbNET COMMUNICATIONS SYSTEM

1.1 Introduction

An earlier JPL Report discussed the continuing interest in distribution automation, the implementation of automation in the electric power distribution system (Kirkham, Friend, Jackson and Johnston, 1990). The primary technical problem facing the implementation of distribution automation, limited data rate on the communications channel, is readily solved by using fiber optics. However, as shown by Kirkham *et al.* in 1989, a rather unusual communications system must be used. The solution described in the 1990 report was a hardware and software communications system designed specifically for distribution automation.

The hardware comprised three major subsystems: a digital fiber optics backbone network installed on the distribution feeders; local power line carrier communications; and a parallel PCM voice channel. The fiber optics backbone operated at medium speed (10 Mb/s), which would provide adequate capability for all foreseeable applications. The PLC system was for local communications to individual customers or loads, and operated at much lower speed. The separate voice channel would be useful for installation, commissioning, operation and maintenance.

In a general-purpose distribution automation scheme, the topology of the power system fixes the topology of the communications network, which can therefore be expected to include a large number of branch points, tap points and interconnections. These features make this communications network unlike any other.

The network operating software has to solve the problem of communicating to all the nodes of a very complex network, in as reliable a way as possible even if the network is damaged, and it has to do so with minimum transmission delays and at minimum cost. The solution we adopted and described in 1990 was called AbNET.

The design of the network and the method of operating it are quite unlike the approach that the telephone company would use, because the system was designed specifically for distribution automation. In the world of the phone company fiber optics are used for trunks, where the design goals are to maximize the distance between repeaters, and to maximize the number of phone calls that can be squeezed into one fiber. This enables the phone company to hold down costs, and to spread the charges among a maximum number of revenue-earning calls. In the distribution automation scheme, AbNET uses repeaters wherever there is a need to tap or branch a fiber, so the distance between them is fixed by the power system and not by considerations of maximizing the use of the resource. The amount of traffic in the network is low by the standards of fiber optics, so there is nothing to be gained by using a high data rate approach. Instead, AbNET maximizes the capability of the lowest-cost system. In terms of the revenue-earning efficiency of the phone company, AbNET is poor. However, its price and performance cannot be beaten in the application for which it was designed.

The selection of operating protocols was modelled after the seven-layer Open System Interconnection hierarchy of the International Standards Organization (Zimmerman, 1980). The choices that define AbNET have an interesting effect: the communications

functions are maximally decentralized and the distribution automation functions are centralized; the two are almost completely segregated. Each aspect can therefore be developed separately, and any AbNET hardware should be able to support any distribution automation software. Competition in both areas is therefore possible, ending an era in which distribution automation was characterized by vertical integration.

1.2 A Summary of AbNET

Since it is necessary to communicate with all parts of the power network, the communications network must be co-extensive with the power system.

The data acquisition and control needs of the system mean that it will be necessary to tap into the communications network at a large number of locations. (For even a single substation, this number could be hundreds.) It is also a requirement that communications be possible to all locations even if the configuration of the network changes, for example because of failure of one or more lines or nodes.

Normally, the distribution system is operated radially, with a limited set of open loops designed to provide an alternative way of bringing power to any given location. The fiber can, of course, cross an open power switch. As a result, the fiber optic communication system is arranged not as a conventional ring, star or bus system, but as a series of interconnected loops, with an occasional spur, as shown in Figure 1-1.

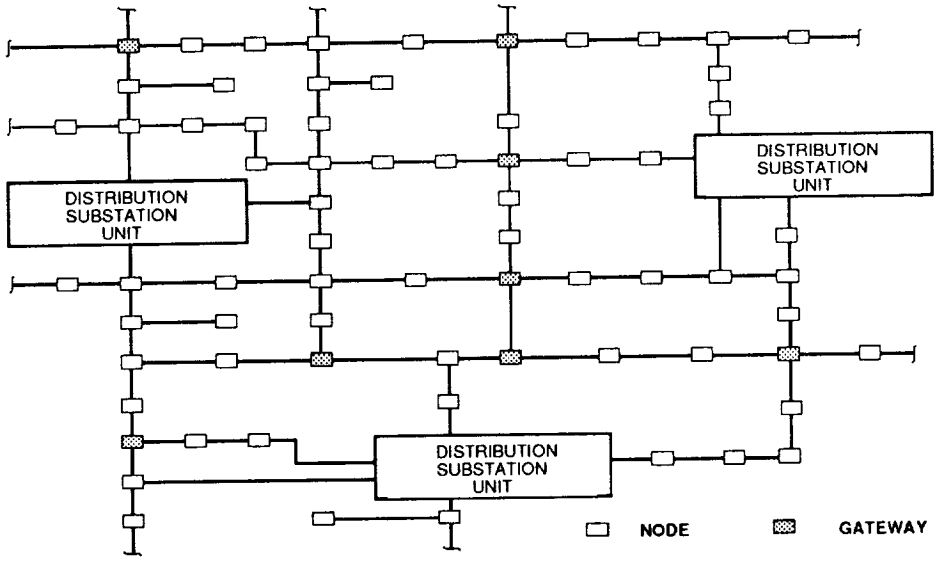


Figure 1-1 Representative topology of distribution system communications

There is no shared medium or common bus. Consequently, most messages will have to pass through many intermediate stations on their way to their destination. The operation of this complex communications system is discussed below in the framework of the Open System Interconnection (OSI) model of the International Standards Organization (ISO). Because the AbNET system was deliberately designed to have an open system specification (ie, use a standard that would be made available to all), it was based on the OSI model.

Computer-to-computer communications in the ISO seven layer model is shown in Figure 1-2. In the Figure, a few words of explanation have been added to the standard

diagram. A more detailed explanation is given in the review below, which is based on the longer description given by Kirkham *et al.*(1990).

The reader who is completely familiar with the OSI model can skip this section: the reader to whom this material is completely new might do better to refer to the earlier report.

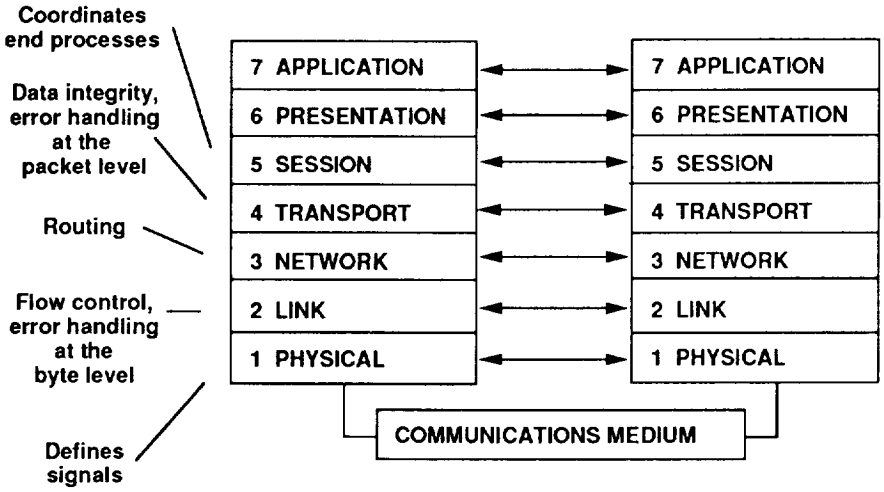


Figure 1-2 Open Systems Interconnection model of ISO

The **application layer** is the layer of software wanting to transfer data across the network. All other layers in the hierarchy exist solely to satisfy the needs of this layer. In the case of a distribution automation system, the applications layer might be the SCADA program, the program that presents data to the operator, or the program that contains the control algorithms which are commonly part of SCADA operations.

The **presentation layer** provides services to the application layer to process the data in some way to make it more suitable for the layers below. This could mean, for example, translation or encryption of the data. In the distribution automation example, this layer could contain user-callable library routines.

The **session layer** is the first of the layers in this hierarchy specifically concerned with communications to another computer. In essence this layer is responsible for coordinating interaction between the opposite end application processes. The layer has to be aware of both the application and the communications. In some applications, the session layer is a "virtual" layer.

The decisions made here are the type of communication to be employed (e.g. full or half duplex) and how the failure of lower layers in the hierarchy is to be handled.

The session layer interfaces to the **transport layer** which is the highest of the seven layers responsible for the integrity of data. For example, this layer would be capable of performing error checking, perhaps on a packet by packet basis.

The job of the transport layer is to furnish error-free messages, in sequence, to the session layer. Whether this is a simple task or a complex one depends on the layers beneath. If the lower layers are retaining message sequence, and are performing error checking, the transport layer becomes very simple. On the other hand, if the lower layers can operate so as to get messages out of sequence, for example, the transport layer must correct the deficiency.

The transport layer connects to the **network layer**, which is responsible for furnishing data to the bare-bit manipulations of the data link layer below.

In some communications systems, the network layer is part of the telephone system. The network layer is responsible for routing the information from node to node. If the telephone system is used, the user has no choice over the route chosen for the data.

The **data link layer** is the highest level at which information as such is handled. The data link layer may perform error control on a word by word basis; for example, parity checking occurs in the data link layer. There may also be some means of flow control or hand-shaking to assure synchronization between devices capable of operating at different speeds.

Some users of the model have divided the data link layer into two sub-layers: **logical link control** and a **medium access control**. The logical link sub-layer is responsible for establishing, maintaining and terminating a logical connection between devices. The medium access control sub-layer ensures that only one device attempts to transmit at a time, performing the function of congestion control. The relevant IEEE standards maintain this distinction.

The **physical layer**, the lowest layer of this hierarchy, is the level at which the electrical or optical signals are exchanged. A specification of the physical layer typically includes a description of electrical and mechanical quantities involved.

For most purposes, and in the case of most standards, it is only the bottom three layers of this hierarchy that are standardized. However, because these layers do need to be standardized for effective communication between different machines, a good deal of work has gone on in this area. IEEE, for example, has developed a series of standards according to IEEE Computer Society Project 802.

These are usually known by the decimal organization of the subcommittees, for example, IEEE 802.3 is CSMA/CD (Carrier Sense Multiple Access with Collision Detection) and IEEE 802.5 is Token Ring.

1.3 Communications Issues for AbNET

In arriving at the AbNET protocols, we made the assumption that the system would have to reach all parts of the distribution system, and be capable of handling all future control and monitoring data. The distribution system topology defined the communications network topology, and the data requirements led us to choose fiber optics as the communications medium. This, in turn, leads to the choice of a baseband system, probably using differential Manchester encoding. The need for multiple taps led to an arrangement of node-repeaters, rather than passive optical power splitters.

The AbNET system is a dual-hybrid network, comprised of a fiber-optics-and-power-line-carrier digital data system in which every node (or RTU) is also a repeater, and a co-located fiber optics PCM voice channel. The nodes are interconnected by fiber optics links routed along the distribution system. The fixed installation of the voice channel is intended to be passive, the optoelectronics and the electronics being carried by the user and inserted into the system as needed. The digital data system, on the other hand, can be put in the framework of the ISO OSI model, as shown in Figure 1-3.

Distribution automation is shown in Figure 1-3 as the applications layer. Descriptions of the various functions that comprise distribution automation will not be repeated

APPLICATION	DISTRIBUTION AUTOMATION
PRESENTATION	ADDRESS TRANSLATION
SESSION	POLLING
TRANSPORT	ERROR CORRECTION CODES
NETWORK	FLOODING ANTIBODY MESSAGE REMOVAL
LINK	SIMILAR TO IEEE 802.5 BUT NO FLOW CONTROL
PHYSICAL	FIBER OPTICS AND POWER LINE CARRIER

Figure 1-3 Main features of AbNET in OSI framework

here. The brief descriptions that follow show how AbNET meets these requirements, and supports distribution automation.

The principal function of the presentation layer is to perform the conversion of data descriptions (in application programs) into addresses or formats that the lower layers can recognize.

The session layer establishes connections between the central unit and the remote nodes on a polling basis. Centralized information about network performance can only be made available to an applications program through the operation of this layer. The session layer therefore retains the ability to command the remote units to establish (and test) connections between adjacent nodes.

The transport layer is centralized at the distribution substation. It normally operates through the local network in support of the data acquisition and control functions. Under unusual circumstances, it can choose to use the local network for access to the area served by another substation. Software error correction can also be performed at this level of the hierarchy, assuming that the lower layers pass corrupt packets into the node.

The network layer is a decentralized message-based, store-and-forward system. Congestion in the network is controlled by an antibody-like algorithm, that allows a node to repeat a message once and only once. Routing is accomplished by network flooding.

The data link layer could incorporate any convenient field descriptions and framing definitions. There is no hand-shaking for flow control, and no request for retransmission in the event of an error.

The physical layer is a fiber optics based hybrid. Multimode fiber will be used, with transmitting optics based on LEDs operating in the near infrared. Cost considerations lead to the choice of a bit rate of about 10 Mb/s, and differential Manchester coding which, because it has two transitions per bit, produces no dc level in the signal. Access to locations on the power line secondary will be by means of low-power high-frequency power line carrier communications.

It will be recognized that the data link layer and some aspects of the physical layer are very similar to some of the specifications in IEEE 802. The AbNET specifications are compared with other communication systems in the next section.

1.4 AbNET Compared to Other Standards

There are several computer communications systems in existence that could have been chosen as a starting point for AbNET. Alternatively, it would have been possible for the

details of AbNET to have been designed from scratch.

Some of the functional requirements for communications in the distribution system can be met by more or less any of the existing Local Area Network (LAN) systems (Hopper, 1986). This is particularly true of the lowest layers of the hierarchy, the physical layer and the data link layer. There are three systems in particular that make good candidates: IEEE 802.3 (Ethernet), IEEE 802.5 and Fiber Distributed Data Interface (FDDI). It was decided that the development time and the technical risk would be reduced by adopting existing standards as far as possible. Therefore, we compared AbNET, 802.3, 802.5 and FDDI.

IEEE 802.3 is a LAN system in which all users are essentially in parallel on the channel, and have equal access to it. If a node wishes to transmit, it first listens to be sure the channel is not already in use (Carrier Sense) and then transmits. Any node can transmit (Multiple Access). If two nodes simultaneously decide to transmit, the event is called a "collision," and it results in the transmitting nodes ending their transmissions, and staying off the channel for a randomly-determined time (Collision Detection). As discussed below, the standard is an IEEE embodiment of Ethernet.

IEEE 802.3 and 802.5 were issued as IEEE standards by the same committee, and it is to be expected that they have some similarities. In fact, they are similar in the data link layer. IEEE 802.5 is a low-end LAN system based on a token-ring protocol. The 802.5 system is based on one originally adopted by IBM as a twisted-pair network operating at 4 Mb/s, when Ethernet was being developed by Xerox using coaxial cable and 10 Mb/s. Both 802.3 and 802.5 use Manchester encoding, and both send source and destination addresses in the packet, as 2 or 6 octets. Both use a cyclic redundancy check based on a generator polynomial of degree 32.

The 802.5 system could itself be regarded as the starting point for the more recent FDDI, the Fiber Distributed Data Interface standard. Whereas 802.5 (and 802.3) use differential Manchester encoding, FDDI uses 4B/5B encoding.* Rather than avoid collisions by sensing a carrier on the channel, as in 802.3, the 802.5 approach was to *prevent* collisions by permitting only the owner of a token to access the medium. This approach was retained in FDDI, although the token transfer process is somewhat different. FDDI, properly known as ANSI X3T9.5 Fiber Distributed Data Interface, is a high-end network designed for fiber optics. Much of the topology, protocols and framing formats used in IEEE 802.5 were retained in the faster network, but some of the traffic-handling is different in order to achieve the 100 MHz data rate of FDDI.

The 802.5 system has one active-monitor, that controls the clocking for the entire ring. This function is decentralized in FDDI, each station being able to transmit at its own local frequency. AbNET is similar to FDDI in this respect, since it is only necessary to design for clock variations between the ends of a two-station link.

Token transfer in an 802.5 node is performed after a transmitted packet is returned. In a large ring this can be inefficient. In FDDI, the token is relinquished immediately after transmission. Tokens are not used in 802.3, and *cannot* be used in a complex interconnected

* In Manchester encoding each bit of data is encoded as a two-bit symbol, essentially a 50% efficient process. 4B/5B encoding transmits 4 bits of data as a five-bit symbol, increasing the coding efficiency to 80%.

system such as required by AbNET.

Traffic control in token-passing rings is accomplished by modifying a parameter in the transmitted packet that controls access to the channel. Both 802.5 and FDDI do this, in slightly different ways. 802.3 has no corresponding function.

AbNET can control traffic from the central location, since only the central station can confer the right to transmit.

1.5 AbNET compared to IEEE 802.3

The best starting place for AbNET development was evidently one of the varieties of IEEE 802. At present, 802.3 or Ethernet-like systems are among the most commonly used methods of connecting computers, including workstations, mainframes and personal computers. Ethernet can be implemented in a number of ways, with some differences in performance. The popularity of the system, and the potential availability of hardware and software to support its implementation warranted closer investigation of Ethernet. Ethernet was developed at Xerox PARC (Palo Alto Research Center) primarily by Robert Metcalfe (Metcalfe and Boggs, 1976). In some ways it can be thought of as a combination of ARPAnet, a store-and-forward packet switching network, and the Aloha network, a radio-based packet-switching system.

The principal variations on Ethernet are Cheapernet and Star-LAN. The IEEE designation uses the following algorithm:

$$< \text{data rate in Mb/s} >< \text{medium type} >< \frac{\text{maximum segment length in m}}{100} >$$

Using this approach, ordinary Ethernet is designated 10BASE5, since it operates at 10 Mb/s, uses a baseband medium (typically a double shielded coax cable, 0.4 inches in diameter) and has a maximum segment length of 500 m.

A "thin" version of Ethernet is designated 10BASE2. "Cheapernet" is another popular name for the standard. It operates at distances up to 185 m over inexpensive RG-58 coax cable.

Star-LAN is designated 1BASE5. It operates over unshielded twisted pair on a point-to-point basis. Ordinarily, the physical medium is the building telephone wiring.

One manufacturer of integrated circuits, National Semiconductor, has produced a series of VLSI devices to implement the functions of IEEE 802.3 for these varieties of Ethernet. There are three circuits in the range. As will be shown below, this arrangement might provide a particularly straightforward way to implement AbNET too.

The chip set is shown in Figure 1-4.

The DP8390 Network Interface Controller connects to the bus of the host computer. It performs a number of functions, which are compared to the AbNET requirements in Table 1-1.

Two of the differences between the Ethernet requirements and the AbNET requirements arise because of the possibility of collisions on the Ethernet system. Since collisions on the medium are not possible, AbNET does not require a jam pattern in the transmit mode, or random time backoff in the event of a collision.

The other difference arises from the fact that Ethernet is not a store-and-forward system. If a node detects a message whose destination address, which precedes the data,

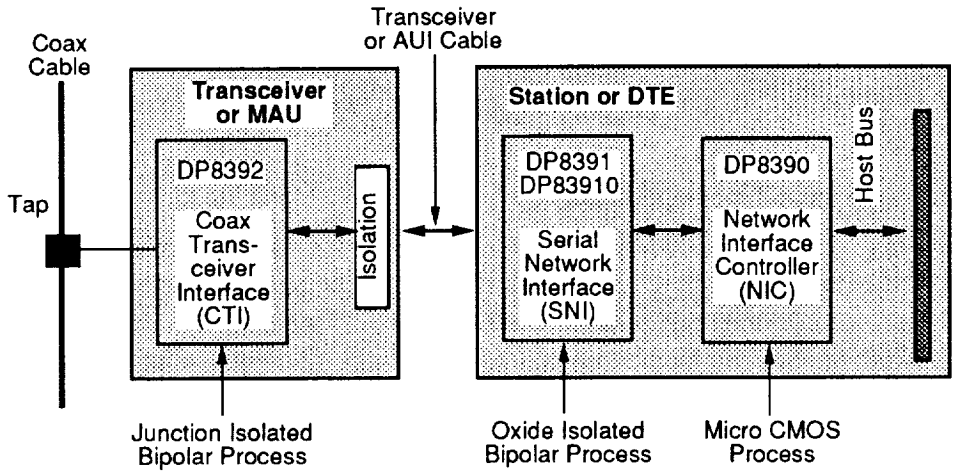


Figure 1-4 National Semiconductor IEEE 802.3 chip set

does not match its own, it can ignore the remainder of the message. Since AbNET is store-and-forward, each node must be able to receive all packets, regardless of the destination address. Fortunately, the Network Interface Controller has a "promiscuous" mode that will allow reception of data intended for other addresses.

The next integrated circuit in the set is the DP8391 Serial Network Interface. This circuit provides the necessary Manchester encoding and decoding, and a number of other functions, as shown in Table 1-2, where again the functions are compared to the AbNET requirements.

The main difference between the Ethernet functions and the AbNET requirements is the provision in Ethernet of differential input and output. In ordinary circumstances, differential I/O is used to provide noise immunity, for example in line driving. In an implementation of IEEE 802.3, it is necessary to drive a length of transmission line between the host computer and the actual network medium, which is normally a coaxial cable. It is possible for there to be noise on the medium, or at the host computer, that appears as a common-mode signal. Such common-mode interference can be rendered harmless with some degree of isolation in the transceiver interface (the third chip in the set implements this), and with differential connections between the transceiver interface and the serial network interface.

In the case of AbNET, the medium is fiber optic, so the connection need provide no further isolation. It is nevertheless more effective at this stage in our system development to use an existing circuit such as the DP8391 and accommodate its differential I/O than it would be to design a new circuit.

The third chip in the National Semiconductor set performs the interface to the actual medium, the coaxial cable. Since the AbNET system uses fiber, this part of the implementation *has* to be redesigned. The details of this are described below.

Since so much of Ethernet seems likely to be repeated in our design, it might be fruitful at this point to clarify what is different about AbNET. Metcalfe himself has observed that the crux of Ethernet was to eliminate the collision problem, ie the simultaneous attempt

Table 1-1. Comparison of IEEE 802.3 and AbNET
implemented using DP8390 Network Interface Controller

IEEE 802.3	AbNET
Receive deserialize	
Allows incoming bits to be clocked into shift register. Identifies Start of Frame Delimiter (SFD) and establishes byte boundaries.	Same as IEEE 802.3
CRC Generate/check	
During Transmit, generates CRC after synch byte, and appends it after last data. During receive, generates local CRC from data, and compares to incoming CRC.	Same as IEEE 802.3
Transmit serialize	
Takes data from FIFO buffer and serializes it for transmission. Adds preamble and synch, appends CRC. In the event of collision, adds jam pattern.	Similar to IEEE 802.3, but no jam pattern required.
Address recognition	
Compares destination address (first 6 bytes of received packet) to physical address. Rejects packet if no match. Can operate in multicast mode.	Cannot be used directly in a system where all nodes are repeaters. A "promiscuous" mode that allows all packets to be received can be selected.
FIFO and FIFO control logic	
This is a 16-byte buffer, with byte counter to start DMA operation. Used in receive and transmit.	Same as IEEE 802.3
Protocol PLA	
Executes collision recovery and random time backoff, formats packets in transmit. Strips preamble and synch in receive.	No need for collision recovery or backoff. Other functions can be used.

to transmit by more than one node (Cox, 1990). In a system that has a common medium, a number of methods of solving the problem have been tried. Ethernet is one; it solves the problem by sensing the presence of a carrier on the medium, and prohibiting transmission

Table 1-2. Comparison of IEEE 802.3 and AbNET implemented using DP8391A Serial Network Interface

IEEE 802.3	AbNET
Manchester encoder and differential output driver	
Accepts data from DP8390 controller, and performs Manchester encoding. Transmits differentially to transceiver.	Same as IEEE 802.3, but no requirement for differential output, since no requirement for isolation of the output.
Manchester decoder	
Receives Manchester-encoded data from the transceiver, and sends NRZ data and clock pulses to controller.	Same as IEEE 802.3.
Collision translator	
Indicates to the controller the presence of a valid 10 MHz signal at its input.	Not used in AbNET
Loopback functions	
Switches input to encoded data instead of received signal. Can be used to check functions except for driver and receiver.	Can be used in AbNET.

until the medium is clear. A collision is a more or less unlikely event, depending on how full the channel is. If a collision is detected, the offending nodes are permitted to retry after a random time. If collisions are rare, then repeated collisions are extremely rare.

The other approach to the collision problem, collision *prevention* by the use of a token as in IEEE 802.5 and FDDI, essentially grants permission to access the medium. This approach has a counterpart in certain single-line railway systems, where the token is a physical entity. As far as distribution automation is concerned, this technique has to be ruled out because it is necessary for each node to pass the token on to its neighbor, and in an interconnected system tokens can be duplicated.

There is no common medium in the AbNET system, so that collisions (in the sense of two nodes attempting simultaneous access to the channel) are not possible. In the flooding AbNET system, a message might arrive at a node simultaneously from two different directions—indeed, this is likely to occur frequently—but provided each node has sufficient buffer space available, no problem is anticipated.

Ethernet is primarily aimed at solving the collision problem. Its adoption on a wide scale has accomplished more: convenient and inexpensive integrated circuits are available

to implement the data link layer of IEEE 802. All sorts of design details, such as packet framing, address format, CRC implementation, Manchester encoding and decoding are already implemented. Clearly, these aspects of Ethernet can be retained in an AbNET system.

The next task is to convert a chip set designed to implement the 802 standards over a wire medium to the AbNET standards over a fiber medium.

1.6 Development approach

The three-chip set from National Semiconductor might form a convenient development system for AbNET. Two of the circuits, the Network Interface Controller that connects to the computer backplane and the Serial Network Interface that connects to the transceiver, are available on an evaluation board.

Although AbNET uses baseband signalling, the transceiver functions have to be electro-optic rather than coaxial, and the third chip in the set, the Transceiver Interface, cannot be used.

Nevertheless, the framing, addressing, CRC and encoding aspects of the chip set make its use in our prototyping work very attractive. To emphasize this point, the problems solved for us by this approach are summarized below.

1.6.1 Packet format

A standard IEEE 802.3 format consists of the field shown in Figure 1-5.

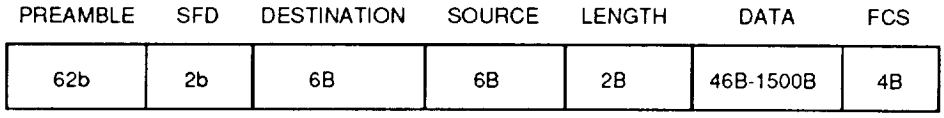


Figure 1-5 IEEE 802.3 Packet format

The **preamble** field consists of 62 bits of alternating ones and zeros. This preamble is used by the hardware to synchronize with the incoming signal (the National Semiconductor chips use a phase-locked loop) and contains no useful information.

The hardware having gained bit synchronization with the preamble, the **start of frame delimiter** is used for byte synchronization. From this point forward in the packet, the receiver knows where byte boundaries are. Both the bit and byte synchronization features of the National Semiconductor chips can be used without change in AbNET.

The **destination** address is normally used to prevent unwanted packages from reaching a node. There are three types of address: physical, multicast and broadcast. The physical address is the address of the individual node. All physical addresses have a MSB of "0". Multicast addresses begin with a MSB of "1". A message with a destination address consisting of all "1"s is a broadcast message, intended for all nodes. As a receiver,

a node can be in a "promiscuous mode" that is capable of receiving all packets. AbNET nodes are expected to act as repeaters, so that they must accept all incoming packets. The broadcast transmission mode and the promiscuous reception mode will accomplish the required lack of selectivity.

The **source** address is the physical address of the originating node. This is of relevance to the central unit in the AbNET arrangement, and is suitably handled by Ethernet hardware. The field is simply passed into the node buffer memory.

The **length** field is a 2-byte field indicating the length of the upcoming data field.

The **data** field contains the message, or a part of the message, being transmitted. It must be at least 46 bytes long, and cannot be more than 1500 bytes long. Messages shorter than 46 bytes can be padded up to the required minimum, in which case the number of valid data bytes is indicated in the length field.

The **frame check sequence** is a 32-bit Cyclic Redundancy Check (CRC) field that is calculated as the packet is being assembled for transmission, and is used at the receiver to check for errors. A standard polynomial known as AUTODYN II is used:

$$X^{32} + X^{26} + X^{23} + X^{22} + X^{16} + X^{12} + X^{11} + X^{10} + X^8 + X^7 + X^5 + X^4 + X^2 + X + 1$$

Packets with improper CRC are rejected.

The convenience of using existing standardized hardware is clear. At the transmitter, the node furnishes the data, and the 802.3 hardware adds the preamble, start of frame delimiter and the CRC. At the receiver, the preamble and start of frame delimiter are stripped by the hardware, and the data (including addresses of source and destination, and the CRC), are passed to the node computer.

1.6.2 Board description

The DP839EB evaluation board plugs into the backplane of an ordinary PC, and provides outputs for Ethernet, Cheapernet and (with an additional board) StarLAN. Figure 1-6 shows a block diagram of the board.

The Network Interface Controller (NIC) handles interactions with the PC, transferring data to and from its local bus. The Serial Network Interface (SNI) performs the Link layer functions, such as Manchester encoding and decoding. It also couples data from the computer, via the NIC to the network, either through a Coax Transceiver Interface (CTI) on the board, or to an external Ethernet transceiver, or to the optional StarLAN daughter board. Our interest is in using the interface to the external Ethernet transceiver to drive an external fiber optics transceiver, of our own design.

1.7 Future work

Because so much of the task is already solved, it is proposed to develop an AbNET implementation based on the National Semiconductor evaluation board (with their concurrence). Our own fiber optic transceiver will be built, suitable for interconnecting the nodes so that the AbNET protocols can be demonstrated in the laboratory using PCs.

This work will be described in a future report.

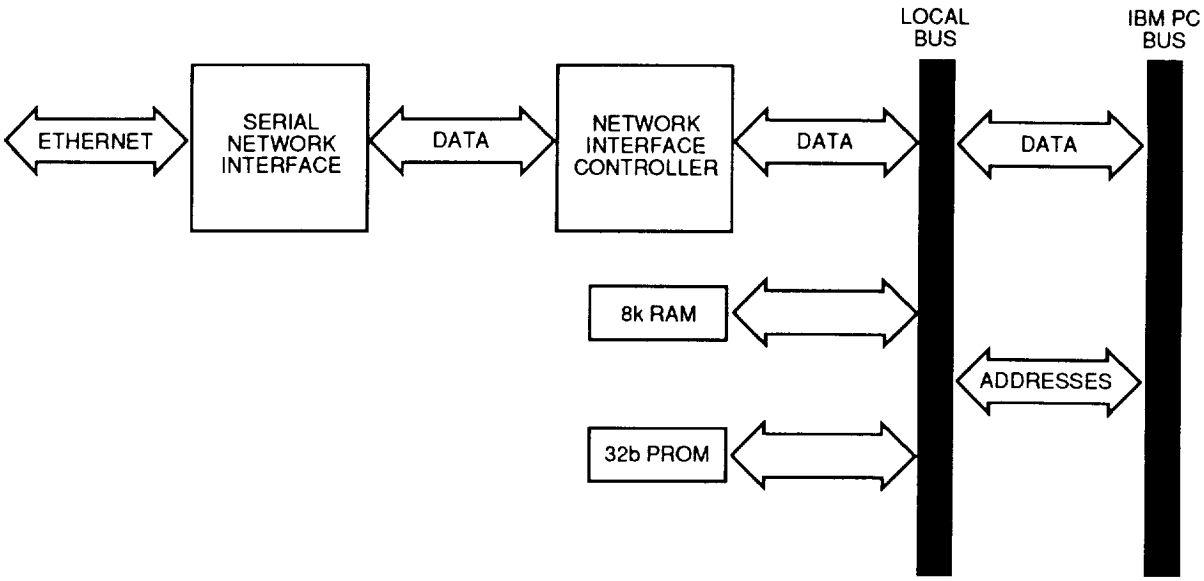


Figure 1-6 Block diagram of evaluation board

REFERENCES

Cox, J. (1990) "Ethernet inventor Metcalfe roots for open standards" Digital News, October 29, p. 66.

Hopper, A., Temple, S. and Williamson, R.C. (1986) Local Area Network Design, Workingham, England and Reading, MA: Addison-Wesley.

Kirkham, H., Friend, H.A. Jackson, S.P. and Johnston, A.R. (1990), Developments in Distribution Automation Applications of Fiber Optics, JPL Publication 90-2, Jet Propulsion Laboratory, Pasadena, California.

Kirkham, H., Johnston, A.R. and Friend, H.A. (1989), Distribution Automation Applications of Fiber Optics, JPL Publication 89-10, Jet Propulsion Laboratory, Pasadena, California.

Metcalfe, R.M. and Boggs, D.R., (1976) "Ethernet:distributed packet switching for local computer networks," Comm. Assoc. for Computing Machinery, (Vol. 19 No. 7), pp. 395–403.

Zimmermann, H., (1980) "OSI Reference Model - the ISO Model of Architecture for Open Systems Interconnection," IEEE Transactions on Communications, Volume COM-28, Number 4, April, pp. 425-432.

SECTION 2.

VOLTAGE SENSOR USING ELECTRO-OPTIC POLYMER

2.1 Background

A considerable amount of work has been done by a number of manufacturers on the measurement of current by optical means, based usually on the Faraday effect. The reason for this is clear. Mathematically, current is the integral of the magnetic field surrounding a conductor, and the measurement device must therefore surround the conductor. The purpose of an instrument transformer (such as a current transformer) is to isolate the information obtained from the high voltage in the system. Any measurement of current must therefore include insulation from the current-carrying conductor. In the case of a conventional current transformer (CT), the insulation must be provided between the current-carrying conductor (the primary winding) and the transformer secondary. If the measurement is optical, the necessary insulation can be provided along the optical path between the sensor and the ground, at a cost significantly less than the cost of insulating a conventional CT. It is therefore possible, in principle, to build an inexpensive optical CT.

Less attention has been given to the measurement of voltage. The reason for this is clear, too. The inherent cost advantage does not apply. Mathematically, voltage is the line integral of the electric field between two points. For a measurement of voltage, therefore, a connection must be made to two points (usually the high voltage line and ground).

It is conventional practice to connect some kind of voltage divider to the high voltage to be measured. This could be a capacitor (of a few nanofarads), a transformer, or (exceptionally) a resistor. In any case, the voltage divider has to be able to withstand the full voltage of the system, and has to be rated for the surges that might be expected. This component therefore dominates the cost of the system. There would seem to be little to be gained by reducing the cost of the sensor device at the bottom of a costly capacitor stack.

There could be a benefit from developing a device for measurement in the low-voltage system (up to about 20 kV) however. If a sensor could be made that could be used directly at these voltages, without the need for a voltage divider, such a device could reduce the cost of obtaining data about the performance of the low-voltage part of the system. At present, there are two obstacles to obtaining such data: the lack of a communication system and the cost of the measurements. Section 1 has described a fiber optic solution to the first of these problems. Section 2 will describe work done to address the second.

2.1.1 Low cost

Electro-optic polymer in bulk form can be produced in large quantities at low cost. However, natural electro-optic crystals will remain costly as crystal growth, location of the crystal axes and polishing are all very complicated tasks. The cost of optical fibers has become quite low, as have costs for LEDs and PIN photo diodes for fiber optic use. Driver electronics, which are mostly operational amplifiers and standard electronic components are typically low cost products. Traditional optical polarization components are expensive. However, for wavelengths in the visible range, it is possible to buy extremely low cost polarizers and 1/4 wave plates made of polymer (Polaroid makes such a product). Many different types and qualities of polarizers and wave plates exist, over a wide price range. If electro-optic polymer could be made transparent in the visible wavelength range, the cost of optical polarization components could come down very much. In recent years, new polarizers and wave plates have come on the market for the wavelength 820 nm. These new products seem to be of good quality, yet are reasonably priced. As demand for these components increases, prices should drop even further. Sensor accuracy depends very much on the optical polarization components, and stringent accuracy requirements may keep the prices of these components high.

Traditional voltage transformers and voltage dividers are very expensive when designed for high voltage levels. The high price is mostly because of high demands on the electrical insulation system at higher voltage levels. In the case of the fiber optic voltage sensor based on an electro-optic polymer, the electrical insulation between the high voltage and the electronics which monitors the signal is less of a problem, because optical fiber is a natural insulator.

A sensor designed to be installed with one electrode attached to ground and the other to high voltage potential would, of course, have to be designed to withstand a voltage higher than the peak phase voltage. However, as the sensor is sensitive to electric fields, it might be possible to operate it without direct electrical contact to ground. If the sensor is mounted between the phase conductor and a surrounding cylinder at a floating potential, the measurement would become sensitive to changes in ambient electric fields. This will probably not be a problem if high accuracy is not a demand. The demands on the insulating system of such a sensor would be much less strict than for the traditional voltage transformers and dividers. As a result, the cost of the insulating system will be greatly reduced and the consequences of a sensor breaking down would be much less severe. Clearly, fiber optic voltage sensors of this kind, based on an electro-optic polymer, can be produced at low enough cost to be feasible for distribution automation.

2.1.2 Small dimensions

The optical components in the head of a voltage sensor based on electro-optic polymer could be made as small as a hazelnut. Even with an electrical insulating system, the unit could be fairly small. The electronics, light source and detector could be located on a standard, rack-mounted printed circuit board.

2.1.3 High reliability and long lifetime

The sensor does not contain any mechanical or electronic parts which are subject to wear. The resistivity of the electro-optic epoxy is so high that virtually no current is able to flow, and the amount of optical power absorbed by the polymer is very low. This means that almost no heat is generated. As a result, there is no wear in the optical system. The only sources of error which would have to be considered are potential changes in the optical polarization components with age, temperature and exposure to electric fields. Light sources such as LEDs have a finite life time. However, the light source does not have to be turned on all the time, only when the sensor is interrogated. Since the LED is not at high voltage potential, it can easily be changed in a routine check.

2.1.4 Low power output

The electric driver circuit for the sensor would typically be operated with operational amplifiers. The output from these can easily be converted to digital form by an A/D converter and be transmitted via a digital communication system.

2.1.5 High immunity to EMI

The optical fibers are inherently immune to EMI. The driver electronics can be mounted in an electrically shielded box, which can be placed in an electrically low noise environment.

2.2. Discussion of Experimental Work

This section describes experimental work on the application of electro-optic polymer to a high voltage sensor. Much of what is discussed is background material. The first four sections 2.2.1–2.2.4 introduce polymer chemistry and fabrication, without going into the deeper theoretical background. Section 2.2.5 explains how the mold and the poling process developed during the experiments. Section 2.2.6 explains the fabrication procedure. In 2.2.7 the test procedure is explained, and finally the test results are discussed in section 2.2.8.

2.2.1 Introduction to the fabrication of EO polymer

Described simply, the fabrication process mixes the two components of an optical epoxy with electro-optic dye, and lets it cure while exposed to a high electric field.

In order to fabricate electro-optic polymer, one needs two basic components:

1 **An electro-optic dye** with a high electric dipole moment μ and a high second order nonlinear polarizability β. The figure of merit for the dye is the product $\mu \times \beta$. The dye molecules are able to align in an electric field. This field will, in the following, be referred to as the poling field.

2 **A polymer**, in the monomers of which the dye is soluble. The polymer must have the ability to undergo a transition from a state of high internal molecular mobility to a state of almost negligible mobility with the poling field applied. In the following, these two states will be referred to as the mobile state and the immobile state.

In order to understand the process, a more detailed explanation is necessary. In order to get useful results, one also has to be very careful about certain steps in the process. Several attempts were made, before a successful sample was produced.

2.2.2 Electro-optic dye

The electro-optic dye is able to exhibit electro-optic effect only when the dye molecules have a statistically finite degree of alignment. In order to align the molecules in an electric field, a high dipole moment is necessary. In the following, the combination of dye and monomer or polymer will be referred to as a dye/polymer system, independent of the state of cure.

Selection of electro-optic dye

For these experiments, the red AZO dye Disperse Red 1 (DR1) was selected. The dye has the chemical systematic name 4-(4 nitrophenylazo)N-ethyl,N-2-hydroxyethylaniline. This selection was based on a number of criteria:

Figure of Merit

Since DR1 has a relatively high product of $\mu\beta$, it is suitable as a dye for the demonstration of electro-optical devices.

Solubility

DR1 is soluble in the monomers of the selected polymer up to 5% (weight). This figure is higher than for the other dyes tested.

Transparency

The DR1/epoxy system in the above concentration is fairly transparent in the near infrared. This is important, because a sensor based on DR1 could be driven by an 820-nm source, which is available in both LEDs and semiconductor lasers. Unfortunately, DR1 is heavily absorbing in the visible spectrum. This makes it impossible to operate a sensor with a He-Ne laser at 633 nm or a visible red LED at 650 nm, in spite of the fact that a 650-nm LED would require much less expensive polarization and optoelectronic devices.

Availability

DR1 is commercially available. Though it was not bought in a ready-to-use form, little work was required to purify it.

Documentation

DR1 is a very well documented dye, several researchers have succeeded in poling thin films of DR1 in polymers. Its properties are fairly well known, an advantage when attempting to fabricate samples in bulk form, which appears not to have been done prior to this work.

As part of the experimental work, other dyes which were believed to have comparable $\mu\beta$ products were tested. All were discarded, either because of poor solubility in epoxy, or because they reacted with the epoxy during the curing process, changing the properties. It is most likely though, that the future will bring new dyes with even better properties than DR1. However, these dyes will probably not be commercially available, but will have to be synthesized in the laboratory.

DR1 is, as mentioned earlier, commercially available. In the form the dye was bought from the manufacturer, it is a powder mixed with some inert material. In order to purify DR1, one has to extract it with acetone. When the other fractions have been filtered off, one is left with pure DR1 in acetone. The acetone can easily be evaporated by applying heat, and one is left with the pure DR1 in dry form. The dye forms crystalline structures in this process, and has to be ground into fine particles in order to facilitate solubility in the epoxy.

2.2.3 Polymer

The following is first a brief introduction into the topic of polymers, and how they can be used for fabrication of electro-optic polymer.

Polymerization means that shorter molecules (monomers) are bonded together chemically and incorporated into longer chain molecules (polymers). Polymers are often solid materials. When this process takes place, the material gradually undergoes a transition from a liquid state to a typically solid state. The internal mobility in the solid state is a function of the nature of the polymer and of the temperature. Some polymers are cross-linking, meaning that chemical bonds occur between the individual chains of molecules. These materials are generally much more stable. Some epoxies are examples of cross-linked polymers.

The polymerization process is often referred to as a curing process. The monomer state is referred to as pre-cured. As polymerization takes place, the system undergoes a gradual change in which more and more chemical bonds are formed. This process can take place in minutes, hours or days. This state is referred to as partly cured. The time it takes from the initiation of polymerization until the cured state has been reached is referred to as the curing time. The curing time can often be significantly decreased by applying heat. When the sample has fully polymerized, it is referred to as cured.

All polymers have a glass transition temperature T_g, above which the material undergoes a transition from a glassy state to a rubbery state. Some polymers will never reach the fully cured state when cured at temperatures below T_g. Though from a macroscopic point of view a partly cured sample might look identical to a fully cured sample, the internal mobility can be much higher. Consequently, for some materials it is necessary to raise the temperature to T_g or above in the poling process in order to obtain maximum stability. Above T_g, mobility will be drastically increased, leading to accelerated decay of the alignment of dye molecules, and the electro-optic effect in the case of an electro-optic polymer.

Therefore, for sensor applications, operation below T_g is necessary, and this is why a polymer with high T_g is desirable. Polymers with high T_g normally have less mobility below that temperature than those with low T_g. Thus, a higher T_g will give a more stable material at lower temperatures. One can ascribe a T_g also to a polymer in its partly cured state. Here T_g is a function of the degree of cure, with increasing values of T_g, the closer the state is to the fully cured state. A general rule is that the higher the degree of cross-linking in a polymer, the higher T_g. Epoxies are examples of polymers with high T_g.

Incorporating an electro-optic dye into a polymer can in principle be done in three ways:

2.2.3.1 Guest/host system

The dye is dissolved in the monomers of the material. When polymerization takes place, a web of chain molecules is formed, in the vacancies of which the dye molecules are trapped and immobilized. During the curing as the web is growing, the vacancies grow smaller and smaller, until the dye molecules are so well confined, that they are not able to rotate.

2.2.3.2 Copolymer system

The dye molecules themselves are attached to the monomers. When polymerization takes place, the dye molecules will be attached to the web of polymerized chain molecules. This is often referred to as the dye being attached to the backbone of the polymer. As the dye molecules are not just confined to the vacancies, but are actually tied to the backbone, this is a more stable system.

2.2.3.3 Dye-polymer

The dye-molecules themselves are part of the monomers. As polymerization takes place, the dye molecules will be incorporated into the web of polymerized chain molecules, as part of the backbone. This configuration is as stable as the copolymer configuration.

2.2.3.4 Selection of polymer

For sensor applications, a long lifetime is important. Therefore, a very stable polymer is necessary. This indicates a material with high degree of cross-linking and high T_g. A copolymer system is preferable, but not essential. These demands eliminated using PMMA,*so epoxy was selected instead. Epoxy can be cross-linking and has a relatively high T_g. The simplest approach was to buy a commercially available epoxy and mix it with the dye. This approach would only provide for a guest/host system, but was chosen for simplicity.

Making a cross-linked epoxy type polymer with DR1 attached to the backbone is theoretically possible, but requires some synthesis in the laboratory. A guest/host system based on DR1/epoxy was selected, and this system has been the basis for the work in this report. The properties of the epoxy however, will also affect results.

Optical properties

The epoxy needs to be transparent at the sensor wavelength. Furthermore, it would be convenient if its index of refraction would be close to that of glass, which is the substrate material for the selected polarizers and the waveplate.

Electrical properties

The dielectric strength is important, as it sets a limit to the strength of the poling field. Also the volume resistivity and the dielectric constant are important. A low volume resistivity could be prohibitive for high voltage applications, as an excessively large current running through the material could cause thermal meltdown. A high dielectric constant could make it difficult to interface the material with electrodes. Even a small air gap would have a voltage drop comparable to that across the electro-optic material. This could lead to insufficient poling and erroneous sensor response.

* Polymethylmethacrylate, commonly known as Plexiglas (US) or Perspex (England).

Table 2-1

Properties of EPO-TEK 301-2

Optical properties	
Transmission	98–99% at 310–2500 nm
Index of refraction	1.564
Electrical properties	
Dielectric strength	500 V/mil
Dielectric constant (25° C)	3.1
Volume resistivity	$> 3.5 \times 10^{14}$
Physical properties	
Hardness, Shore D	82
Water absorbtion: (30 days; 94% RH; Room Temp)	0.01%
Tg (cure 80° C, 90 minutes)	65° C
Curing time 80° C	1.5 hours
Curing time 25° C	2 days

Physical properties

A relatively long cure time is necessary in order to allow the dye and the monomers to mix before poling. A T_g higher than the operating temperature of the sensor is necessary. Low viscosity in the precured state is necessary to get air bubbles out. Low water absorption is important in order to avoid increased conductivity with time. A reasonable hardness is also important to make the material practical to work with.

The epoxy selected is EPO-TEK 301-2 manufactured by Epoxy Technology Inc. Some of its more important properties are shown in Table 2-1.
Comments on specifications for EPO-TEK 301-2:

The optical properties of 301-2 are excellent for application to making electro-optic polymer. It has very little attenuation at either 650 nm or 820 nm, so it presents no optical limitations. Note that the index of refraction is expected to increase when DR1 is added. The dielectric strength of 500 V/mil sets a limit to the poling field. The most common units used in describing poling fields are MV/cm. 500 V/mil is equal to 0.196 MV/cm. Poling fields in the order of 10 times this value have been documented for thin film applications utilizing other polymers. An increase by at least a factor of 10 in dielectric strength would be desirable. The dielectric constant and volume resistivity are not limiting factors. Hardness and water absorbtion are not problems. T_g is above expected operation temperature, but still a higher T_g would be desirable in order to increase the safety margin, and to get an even more stable material at lower temperatures.

2.2.4 The poling process

For simplicity, the combination of dye and monomers/polymers will be referred to as the dye/polymer system independent of the state of cure and independent of whether the reference is made to a guest/host system or to a copolymer system. In order to perform the poling process and maintain the electro-optic effect, it is necessary to be able to generate two states in the dye/polymer system, one where the degree of mobility of the dye molecules is high, and one where it is negligible. There are different approaches to this problem depending on the type of polymer being used.

The principle of the poling process is to expose the dye/polymer system to an electric field in order to make the dye molecules line up in the direction of the field. The electric field is referred to as the poling field, and the alignment of the dye molecules is referred to as poling. In order to maintain poling, the dye/polymer system has to undergo the transition from a state of high mobility of the dye molecules to a state of negligible mobility with the poling field maintained. The poling apparatus is shown in Figure 2-1. A schematic of the poling arrangement is given in Figure 2-2.

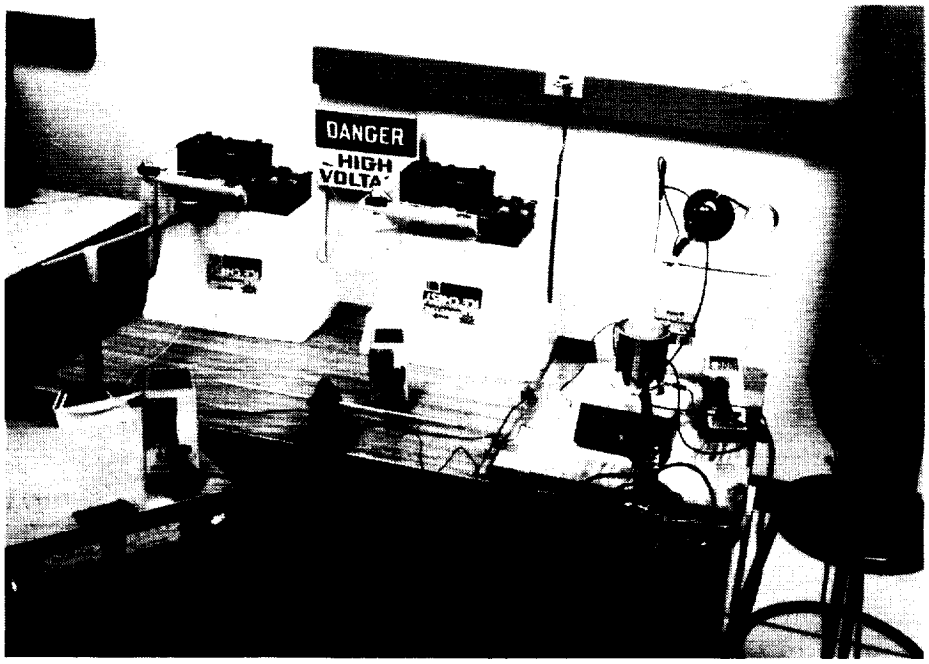


Figure 2-1 The poling apparatus

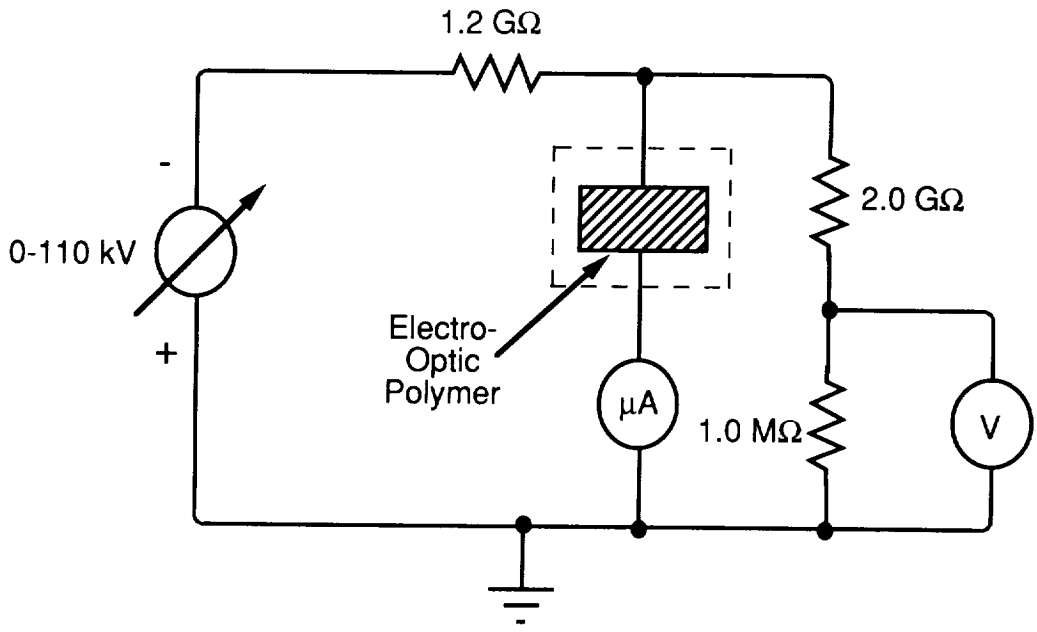


Figure 2-2 Electric schematic of the poling

2.2.4.1 Thermo-poling

In thermo-poling the temperature of the dye/polymer system is raised to above T_g, the poling field is applied, and the system is cooled to below T_g with the poling field maintained. This method has been successfully used to fabricate electro-optic thin films based on the system dye/PMMA, using guest/host systems as well as copolymer systems. The copolymer systems have been the most stable, but unfortunately PMMA is non-cross-linking and has low T_g, which gives it a lifetime too short to be acceptable for sensor applications. Thermo-poling as described above has not been attempted with the DR1/EPO-TEK 301-2 system.

2.2.4.2 Cure-poling

Cure-poling refers to a process in which the poling field is applied to the dye/epoxy system already in a partly cured state, and the field is maintained until the cured state has been reached. This was the first method that was attempted for poling DR1/EPO-TEK 301-2, but with no positive result. The sample was pre-cured for 17 hours before any attempt was made to turn the poling field on. As epoxy is ionic conductive in the precured and partly cured states, the electric field was increased very slowly in order not to damage the sample. At a level of 100 kV/cm it was kept steady. The poling field was maintained for another 12 hours. The sample was polished and ready to measure a few hours later, but no electro-optic effect was detectable. The negative result was interpreted as being caused by incomplete cure of the epoxy. Though the epoxy appeared to be fully cured, the internal mobility was probably too high to maintain any poling. As 301-2 is a very slowly curing epoxy, 100% cure might take weeks to obtain at room temperature, or might never be obtained. To solve this problem, a combination of thermo-poling and cure-poling was tried.

2.2.4.3 Combined thermo- and cure-poling

This process starts with the dye/polymer system in a pre-cured state. The poling field is gradually turned on, and at the same time heat is applied. The current through the sample, which is a control parameter for the process, is monitored, and the poling field is regularly adjusted, using this information. When the temperature reaches a certain predetermined value above T_g, it is held constant. The current is a good indicator of the state of cure. When the current begins to drop appreciably, this indicates that the sample is approaching cure, and it is left for an additional period of time to be absolutely sure that it has been 100% cured. When this time has passed, the sample is left to cool down to room temperature while maintaining the poling field. When room temperature has been reached, the poling field is turned off, and the poling has been completed. This proved to be a successful method.

2.2.5 The mold and electrodes

In order to perform a poling, one must be able to place the dye/polymer system between a couple of electrodes which, with a suitable voltage applied across, can provide for the poling field. Theoretically the dye/polymer system does not have to be in electrical contact with either of the two electrodes, but can be placed somewhere between the electrodes with the rest of the gap filled with air or another dielectric. However, rather high electric field strength is needed, and a gap in series with the dye/epoxy system would cause an extra voltage drop, which would have to be compensated for by a higher voltage. Therefore, it is more convenient to let the dye/polymer system be in direct electric contact with the electrodes. Some kind of containment is necessary to keep the liquid dye/polymer system in place between the electrodes. This is provided for by the mold. The mold has to be made out of a dielectric able to withstand the same high electric fields as the dye/polymer system.

The development of the mold started with a rather simple approach, which was later refined into a more complicated device in order to overcome difficulties in the poling process. In the following, the DR1/EPO-TEK 301-2 system will for the cause of simplicity be referred to as the epoxy.

Figure 2-3 The first mold: made out of a 1/2 in. sheet of teflon, with a hole covered with aluminum tape. A sample can be seen on top.

2.2.5.1 The first mold

The first mold, shown in Figure 2-3, was made simply out of a sheet of Teflon 1/2 in. thick and 12 × 12 in. wide. A 1/2 in. hole was drilled in the center, one side was covered with a strip of conductive aluminum tape, the hole was filled with epoxy, and the other side was sealed with the tape. The mold was sandwiched between a high voltage electrode and a ground potential electrode. Because the surrounding air has a much smaller dielectric strength than epoxy or teflon, it was necessary to give the mold very thick walls in order to avoid breakdown. The wall thickness, which is about 6 in., provides an air gap between the electrodes of about 12 in., enough to withstand more than 100 kV. In this mold the cure poling was attempted but without any success.

The next step was to try the combined thermo- and cure-poling. This was first attempted simply by placing the ground potential electrode on top of a hotplate. The heat would be conducted from the ground potential electrode to the epoxy.

This attempt was not successful, as the non-uniform heating of the mold made it bend. When the epoxy reached its solid state, due to the bending of the mold and the different thermal expansion of the two materials, a crack formed in the interface between the mold and the polymer. The crack then filled with air, which was ionized by the electric field. This resulted in electric arcing, making it necessary to interrupt the poling process.

2.2.5.2 The second mold

In order to avoid the crack formation a new type of mold was developed as shown in Figure 2-4. A 5 mm slice was cut out of a polystyrene cuvette with an inner cross section of 10 × 10 mm. The slice was cemented between two copper electrodes which each would fit the cuvette cross section as a lid. One of the electrodes was tapped with two threaded holes, one for filling in the epoxy, one for letting out the air. The two holes could be sealed with screws after the mold was filled with epoxy.

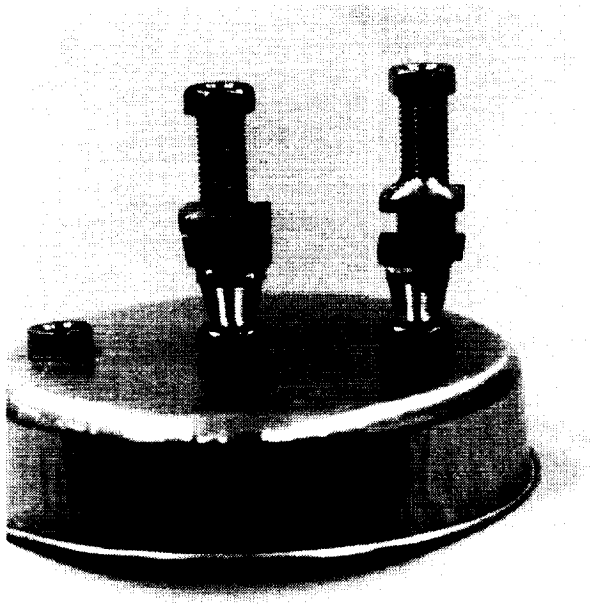


Figure 2-4 Samples fabricated in the second mold. The electrodes remain attached to the sample.

A special apparatus was developed to perform the poling. The thin polystyrene walls of the cuvette would not provide enough insulation distance to avoid break down of air. Consequently it was decided to let the poling take place in insulating oil. The oil would not only provide electrical insulation, but would also work as a medium for heat transfer, providing a very uniform heat to the sample. A special metal container was developed for the purpose. The setup is shown in Figure 2-5. The container can be electrically grounded, and at the same time placed on a hotplate. The container provides electrical connection from the ground potential to the bottom electrode of the sample, without direct thermal contact between the sample and the container. In this way the sample is in thermal equilibrium with the oil. The high voltage is led to the top electrode through a cable insulated by a teflon tube, which is partly submerged in the oil.

The first sample exhibiting electro-optic effect was processed in this type of mold.

However, the mold had some disadvantages. Because of the sharp edges of the electrodes, which led to intensification of the electric field, dielectric breakdown occurred at lower voltage than expected by just taking the spacing between the electrodes into account. Also, the mold was very difficult to fill without leaving vacancies of air behind. Another

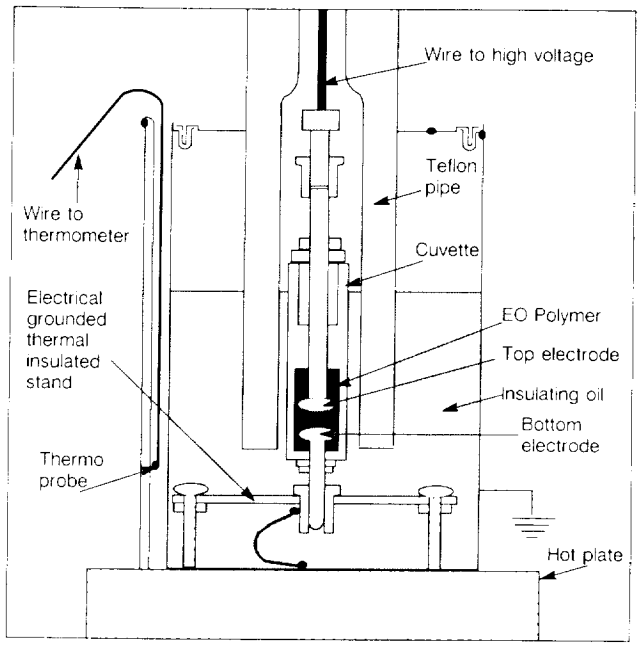


Figure 2-5 Poling chamber filled with electrical insulating oil

problem was that the hot oil attacked the polystyrene walls of the cuvette, making it necessary to remove the polystyrene and grind and polish the epoxy sample later.

2.2.5.3 The third mold

Because of this partly dissatisfactory result, an improved mold was developed. The third mold is depicted in Figure 2-6. It was made out of a glass spectrophotometric cuvette. A hole was drilled in the bottom, and the ground potential electrode was mounted there. The high voltage electrode was mounted to reach into the cuvette from the top. This type of mold was used in the same poling apparatus as previously described. Due to rounded electrodes it was possible to use a much higher poling field, and consequently a significantly higher electro-optic effect was obtained. Because of the glass walls, no grinding or polishing was needed. The sample could be taken directly from the poling apparatus and placed in the optical test set up.

Some cracks had formed in the glass. These were not a problem during the poling, as they were filled up with oil. However they set a limit to the modulation field in the test set up.

In order to avoid this problem, a polystyrene cuvette was used for the later samples. This final version is shown in Figure 2-7. The polystyrene becomes soft and pliant at 80° C, eliminating the problem of crack formation. In order to protect the polystyrene from attack from the hot oil, it is necessary to cover the outer walls of the cuvette with a thin layer of regular "5-minute" epoxy. This, of course, ruins the optical surface quality of the cuvette. However it is possible to break and remove the polystyrene cuvette without damaging the epoxy sample. As the inner surfaces of the cuvette are very smooth, the optical surface quality of the sample, when the cuvette has been removed, is good enough that polishing is not necessary.

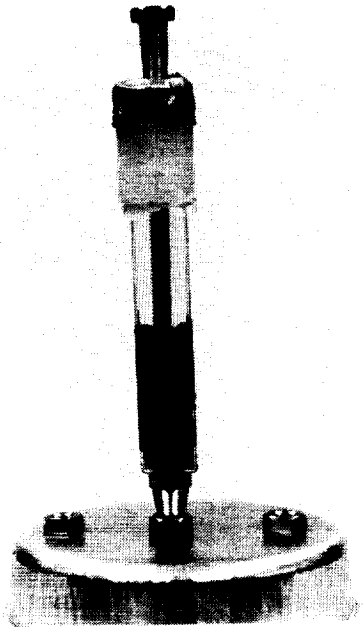


Figure 2-6 Sample fabricated in the third mold. The mold is made out of a glass cuvette, which cracked during the poling process.

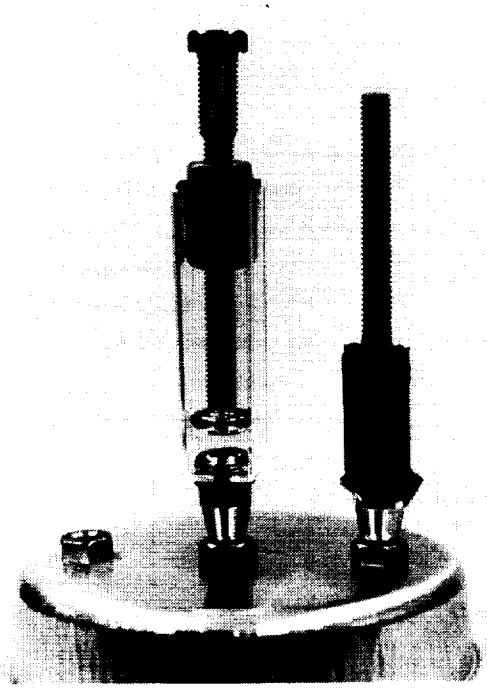


Figure 2-7 The fourth mold. This is the final version for the fabrication of test samples. The mold is here seen before it is filled with the resins of electro-optic polymer. Next to the mold is shown a poled sample after removal of the polystyrene walls.

2.2.6 Fabrication procedure for EO polymer

In the previous section, a broad discussion was given about the materials, the poling process, and the history of the evolution of the poling process and the mold. As such, it served to explain the background of the fabrication process. The purpose of this section is to clarify the concept by recapitulating the fabrication procedure in a concise form.

2.2.6.1 Materials

The electro-optic polymer developed is based on a transparent epoxy as the host polymer, mixed with the electro-optic AZO dye Disperse Red 1 (DR1), systematically known as 4-(4-nitrophenylazo)N-ethyl,N-2-hydroxyethylaniline. As host polymer, the optically transparent epoxy EPO-TEK 301-2 from Epoxy Technology Inc. was selected. DR1 is soluble in the monomers of the epoxy at concentrations up to 5% (by weight) above which saturation takes place.

2.2.6.2 Preparation of the materials

DR1 is purified by extraction with acetone. The acetone is evaporated, and the pure dye, which forms crystals, is ground into a fine powder. The red powder is mixed with the epoxy in its monomer state to 5% concentration. In order to dissolve all the dye, heating and stirring for a long time is necessary. Stirring with a magnetic spin bar for days has given the best results. When all the dye is dissolved, the amine is mixed in, and the cure of the epoxy is initiated. It is important to mix long enough that there is no sign of phase separation between the epoxy and the amine. After this has been accomplished the compound is brought into a vacuum chamber, and vacuum is applied until it is evident that most air bubbles have left the sample.

2.2.6.3 Pre-cure

The epoxy compound is now poured into a mold, with both a bottom and top electrode in direct contact with the epoxy. The mold is placed in a vacuum chamber in order to remove all air bubbles. The sample is left in the vacuum chamber to partly cure for 15–20 hours in vacuum.

2.2.6.4 Poling of the sample

The next step is to place the mold in an oil bath, with the electrodes connected to a 100-kV (dc) high voltage supply. The oil bath works as a medium for heat transfer and electrical insulation. When the heat is turned on, the increased temperature accelerates the curing process. When the sample has reached about 40° C, the poling process can begin. The poling process is the alignment of the dye molecules in the electric poling field. The alignment of the dye molecules provides for the electro-optic effect in the fully cured polymer.

Voltage is gradually applied across the sample, at the same time as it is being heated. The current through the sample is constantly monitored, and the voltage adjusted, in such a manner that the current does not exceed a preset limit of 10μA. When the temperature in the oil bath reaches 80° C, it is kept stable. At 80° C the cure time is only 1 1/2 hours.

As the state of cure proceeds, the current gradually drops as a function of the decrease in ionic conductivity. The voltage can therefore frequently be turned up to keep the preset current value. This procedure is repeated several times until the maximum electric poling field of approximately 180 kV/cm has been reached. Simultaneously the current decreases to its minimum of a few μA or less.

The decreased current indicates that the sample is close to having reached its fully cured state. The temperature and voltage are maintained for a period of time, to make sure that the sample is 100% cured. Considering the pre-cure time, and the long time it takes to heat up the sample to 80° C, 45 minutes to 1 hour should be allowed. Then the heat is turned off, and the system is cooled down to room temperature, still with the poling field applied. When room temperature has been reached, the poling process is completed, and the mold can be removed from the oil bath. Once a sample has been fabricated, the next step is to test whether its electro-optic properties are satisfactory. This is done on an optical bench.

If the sample has been cast in a glass cuvette, it can be taken directly from the oil bath, cleaned and mounted in the optical test system. If the sample has been cast in a polystyrene cuvette, it is necessary to first remove the polystyrene walls. With care, this can be done without any damage to the optical surfaces of the sample. If damage is done, the optical surfaces will have to be polished.

2.2.7 Test of electro-optic properties

The test system is, from an optical functional point of view, identical to the fiber optic sensor system, the Pockels cell, which the electro-optic polymer is going to be an integrated part of. The sensor system is much like a transversely modulated electro-optic laser modulator. However, instead of using fiber optics, discrete optical components on an optical bench are utilized. This system allows greater flexibility. It is much easier to mount the sample, and there are more degrees of freedom in adjusting the polarizing components.

The test system, seen in Figure 2-8, works this way: The light source is a horizontally linearly polarized infrared He-Ne laser (1150 nm, 0.7 mW). In order to give maximum sensitivity and linearity to the system an offset phase retardation of 1/4 wave, which changes the polarization state of the light from linear to circular, has to be provided. This is normally done with a 1/4 wave plate, mounted with its fast axis in an angle of 45 degrees relative to the polarization plane of the incoming light. However, 1/4 wave plates for 1150 nm are not standard components, so a Babinet-Soleil compensator was mounted with its fast axis at the same angle as mentioned for the 1/4 wave plate. The Babinet-Soleil compensator is a variable retarder (waveplate), and it is adjusted to give an offset phase retardation of 1/4 wave. The Babinet-Soleil compensator can be adjusted, unlike a 1/4 wave plate, to compensate for intrinsic retardation (birefringence) in the sample. After having passed the Babinet-Soleil compensator, the beam passes through the sample of electro-optic polymer. The electro-optic polymer, like the 1/4 wave plate, works as a phase retarder. The phase retardation originating from the electro-optic polymer consists of two components. One component is the intrinsic retardation (birefringence), which can be compensated for by the Babinet-Soleil compensator. This component is believed to be insignificant in a practical sensor system even without a compensator.

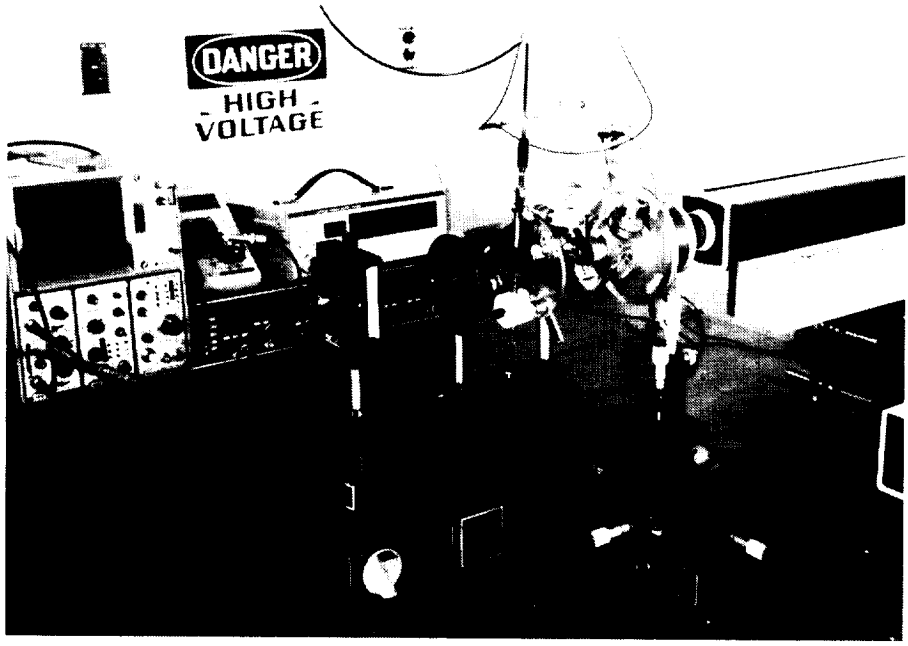


Figure 2-8 The test set up on the optical bench for test of electro-optic polymer

The other component is the phase retardation induced by the voltage across the sample due to the electro-optic effect. This is the component we want to measure. As this component is proportional to the voltage across the sample, it is in this case a 60-Hz waveform, with an amplitude proportional to the amplitude of the voltage.

The sum of the three contributions of phase retardation, two coming from the electro-optic polymer, and one from the Babinet-Soleil compensator result in a net polarization state of the light once it has passed these two optical components.

The two time-invariant components together result in the 1/4 wave phase shift. This phase shift brings the light from the linear polarization state to a circular polarization state. The time varying signal (60 Hz) modulates the radius of the circle at an angle of 45 degrees relative to both the fast axis and the slow axis of the electro-optic polymer. In this way, the circle is modulated between two elliptical states, a tall and slim and a short and fat elliptical state.

No detector is able to analyze the phase of the light directly, so the polarization modulation has to be converted to a modulation in optical power. This is done by an analyzer. The analyzer is a linear polarizer. It is mounted with its polarization axis at an angle of 45 degrees relative to both the fast axis and the slow axis of the electro-optic polymer. This yields four different solutions. In two of those, the resulting modulation in optical power will be in phase with the original electrical modulation signal. In the two others, it will be in anti-phase.

The transmitted optical power as a function of time consists of a relatively large dc part, with an ac ripple on top. The ac ripple is a duplicate of the electrical modulation signal. Though the dc part of the signal does not contain information about the electrical modulation it is still important because it can serve as a reference. The ratio of the ac part relative to the dc part is a measure for the electrical modulation signal, as it is independent

of the injected optical power and the loss in the system.

The transmitted optical power is incident on a Ge photo diode, and measured by an optical power meter. The electrodes mounted on the sample of electro-optic polymer are connected to a high voltage power supply, and the sample is modulated with a 7.96-kV, 60-Hz sinusoidal voltage. The dc level of the optical signal is measured by the power meter, and the ac modulation originating from the electro-optic modulation is monitored on an oscilloscope, and its RMS value measured with a lock-in amplifier synchronized with the voltage supply.

2.2.8 Test results

The following data were recorded from the test of the most successful sample. The sample had been poled in the combined thermo-cure-poling process. It was poled in the fourth type of mold, a polystyrene. The poling was performed on February 20, 1990.

2.2.8.1 Electro-optic effect

The optical power output from the laser was 1.32 mW and the power transmitted was 263 μW. At a modulation voltage of 7.96 kV, the modulation index (the ratio of the RMS value of the ac signal relative to the dc level of the transmitted power) was measured. The value was 4.2×10^{-3}. The electro-optic coefficient for the polymer can be calculated from this result. The value is $r_{33} = 2.2 \times 10^{-14}$. It should be emphasized that the modulation index changes greatly with the position of the laser beam in the electrode gap. This indicates that the poling of the electro-optic polymer is not homogeneous. The above value was the *maximum* that could be obtained, when optimizing the position of the beam in the electrode gap, the adjustment of the polarizers and the 1/4 wave plate.

The value of the electro-optic coefficient is much smaller than for electro-optic crystals. However, for utilizing the electro-optic polymer in a high voltage sensor the plan is to have the voltage, which is in the order of 10–20 kV, directly applied across the material. This results in an electric field large enough to provide an easily detectable signal, even for this small value of electro-optic effect. To improve the signal to noise ratio, it would be worthwhile trying to increase the electro-optic effect by a factor of 10–50. This value is within the range of what is suitable for a high voltage sensor. A much higher electro-optic effect would lead to inaccuracy due to a nonlinear response of the sensor.

2.2.8.2 Absorbance and index of refraction

An absorbance spectrum was recorded for the sample. The spectrum can be seen in Figure 2-9. Numerical values for the absorbance were recorded at two wavelengths, 820 nm and 1100 nm. The sample has its lowest absorbance at 1100 nm. The two values are $A_{(820nm)} = 0.355$ and $A_{(1100nm)} = 0.097$. The index of refraction for the sample was also measured to be in the range $n = 1.585$–1.592. For this report, a value of 1.59 was used as representative.

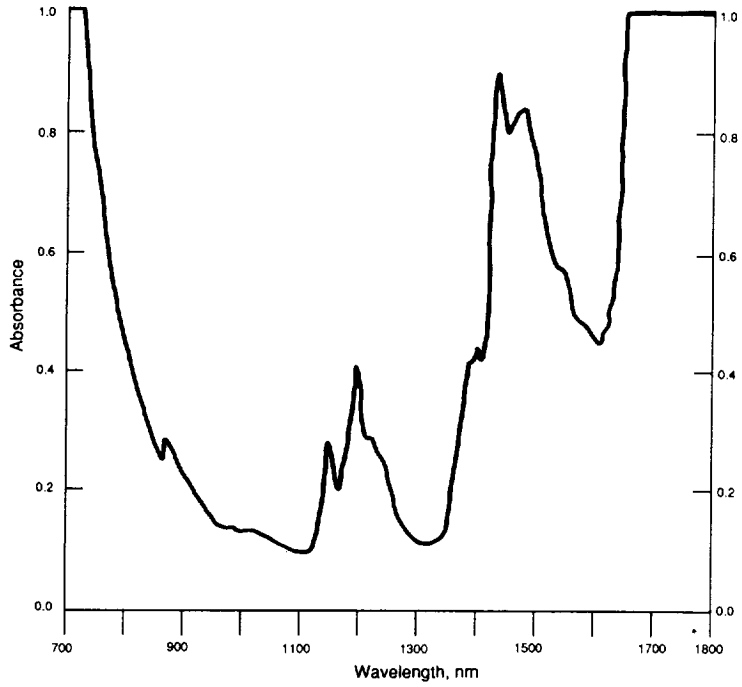


Figure 2-9 Absorbance as a function of wavelength

2.3 Design Approach for a Fiber Optic HV Sensor

The purpose of developing a poling process for fabrication of electro-optic polymer was to demonstrate that this material can be used for voltage measurement at high voltage levels. To demonstrate this, a fiber optic AC voltage measuring system has been built. The system has been built only for demonstration purposes, and development work remains to be done to create a device appropriate for industrial mass production.

2.3.1 Discussion of the design approaches

In order to make the design process as simple as possible, it was decided to build the fiber optic sensor to reflect the laboratory set up on the optical bench as closely as possible. It was therefore decided to use the same polarimetric configuration as described in Appendix A, section 3.1.

2.3.1.1 The fiber optic transmission system

The light is sent through the polarizer, the 1/4 wave plate, the electro-optic polymer sample, and the analyzer in a straight beam, exactly as it was done on the optical bench. One difference between the transmission principle in the fiber optic sensor and that of the optical bench is that in the fiber optic sensor, the straight beam is provided by collimating the beam from an optical fiber with a GRIN rod lens (GRaded INdex rod lens) instead of using the straight laser beam. When the beam has passed the analyzer, it is focused by another GRIN rod lens into the core of an optical fiber. The optical fiber transmits the optical signal to the detector. In the same way, the optical power is brought from the light source, which is an 820 nm LED, to the sensor. The optical polarization components are, in principle, just miniature versions of the same type of components used on the optical bench. The principle of the fiber optic transmission system can be seen in Fig. 2-10.

2.3.1.2 Adjustment of polarization components

In order to position the optical polarization components at the correct relative angles, some means of rotational adjustment is necessary. To provide this, the optical sensor was fabricated in four parts, which can be rotated independently.

The electro-optic polymer with electrodes constitute one part. This part is potted into the sensor housing, which provides for electrical insulation and mounting for the other rotating part.

The 1/4 wave plate is cemented to the end of a short pipe which is able to rotate inside a hole in the sensor housing. Each of the two polarizers are mounted in the same way, cemented to the end of a hollow cylinder, which also contains a GRIN rod lens. The polarizers and the GRIN rod lenses are cemented together with optically transparent epoxy. The other end of the cylinder has been made to mount to a fiber optic SMA connector. When the connector is mounted, the polarizer, the GRIN rod lens and the optical fiber constitute one part, from a mechanical perspective. In the following, the two parts will be denoted the polarizer and the analyzer, dependent on their function in the system.

The analyzer is mounted in a hole in the sensor housing where it can be rotated relative to the electro-optic polymer sample.

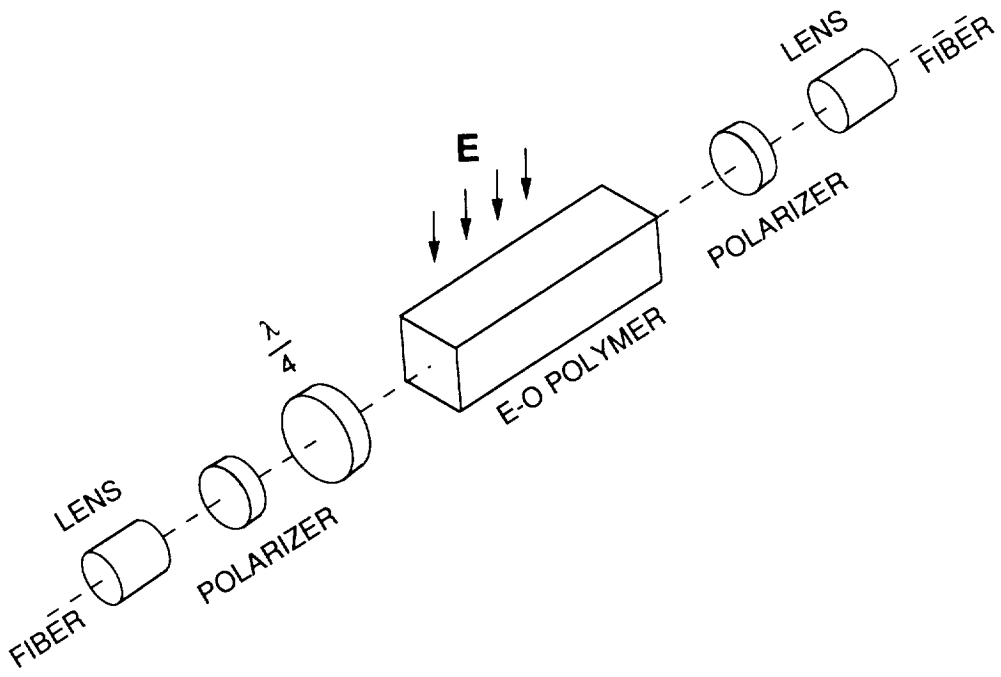


Figure 2-10 Principle of fiber optic transmission system

The polarizer is mounted inside the short pipe which holds the 1/4 wave plate. In this way, the polarizer and the 1/4 wave plate are able to rotate independently, relative to each other, and relative to the electro-optic polymer in the sensor housing.

2.3.1.3 Sensor housing and electrodes

In addition to the optical polarization components and the electro-optic polymer, the sensor consists of the sensor housing and a top and a bottom electrode. The sensor, with its driver electronics, is shown in Figure 2-15.

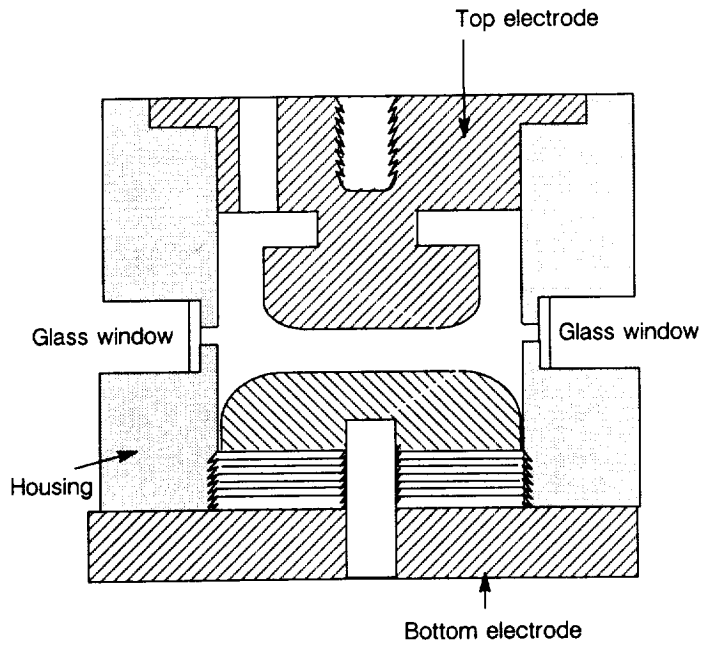


Figure 2-11 The first sensor mold

The first sensor mold

The first approach to the design of the electrodes and the sensor housing can be seen in Figures 2-11 and 2-12.

The idea for the first sensor design was to build the sensor housing and the electrodes as a mold which could hold the constituents of the electro-optic polymer in their liquid state. This mold would then be placed in the poling apparatus immersed in the oil bath. The mold had optical windows mounted in the same holes where the polarization components were meant to be mounted. The windows would keep the electro-optic polymer in place in the liquid state, and still provide access for the light beam. The mold had a copper pipe mounted in the hole in the top electrode. This pipe would make the inside of the mold open to the air, so pressure would be able to escape if it built up during the poling process. At the same time, the pipe would serve as a conductor from the high voltage source to the electrode. The bottom electrode would be grounded, as the mold would be sitting on the bottom of the grounded metal container holding the insulating oil. An attempt was made to pole a sample of polymer in this special mold in the same way that test samples had been poled. This was unsuccessful. At some time in the poling process dielectric breakdown occurred and the polymer was never fully poled.

The mold was taken apart, and it was discovered that a gap had formed between the polymer and the wall of the sensor housing. As the mold was open to the air, air had been sucked in. This air had ionized, and an electric arc had been formed, resulting in further damage.

The formation of the gap is believed to originate from the combination of two causes:

1. Epoxy does not bond very well to Delrin, which is the material of which the sensor housing was made. As a result, there was almost no adhesion between the two surfaces.

2. When epoxy cures, especially at higher temperatures, it shrinks. The larger the volume, the larger the decrease in linear dimensions.

The epoxy sample, which was considerably larger than the test samples, had simply been pulled away from the walls of the housing by its own shrinking, and an air gap had been formed.

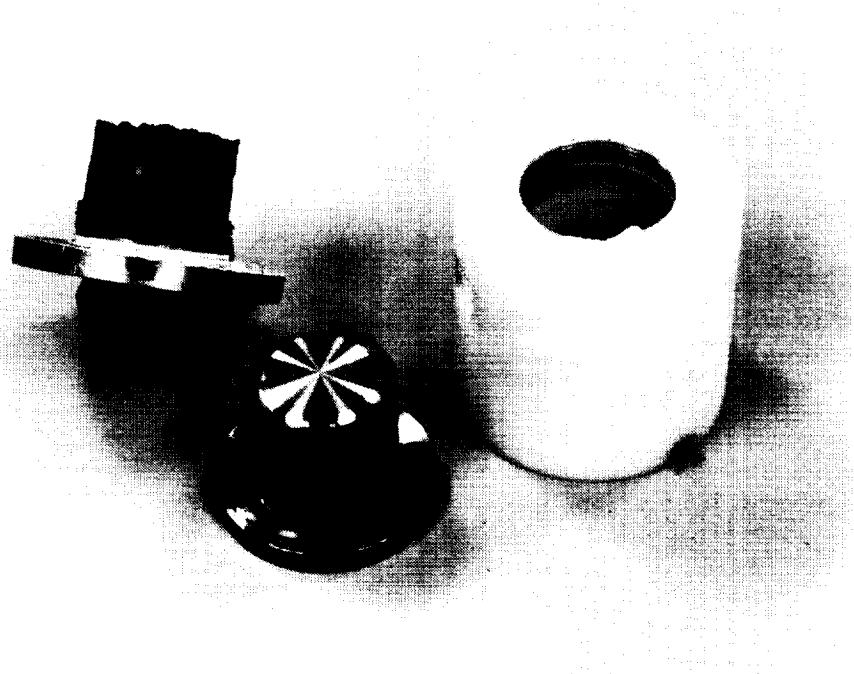


Figure 2-12 The first sensor mold disassembled after poling failure

The second sensor mold

A new approach to the design of the sensor uses an already poled sample as the electro-optic material. This new approach can be seen in Figures 2-13 and 2-14.

The mold has been slightly redesigned. The top electrode has been removed, so the mold is open to the air. The bottom electrode has been redesigned, so an already poled sample can be mounted with its bottom electrode in electrical contact with the bottom electrode of the mold. Except for these alterations, the mold was kept the same. A good sample of electro-optic polymer was mounted in this new mold, and the spacing between the sample and the wall of the mold was filled with the transparent epoxy EPO-TEK 301-2, which has also been used in the fabrication of the electro-optic polymer. The epoxy resins were subjected to vacuum in their liquid state in order to remove all air bubbles. The epoxy was then left to cure at room temperature.

In this sensor design, the poling process and the mounting of the electro-optic polymer in the sensor housing have been separated. This is an advantage if something goes wrong with the poling process. Another advantage of this design is that its electrical insulation properties are independent of the ability of the epoxy to bond to the sensor housing, because the surrounding epoxy provides electrical insulation.

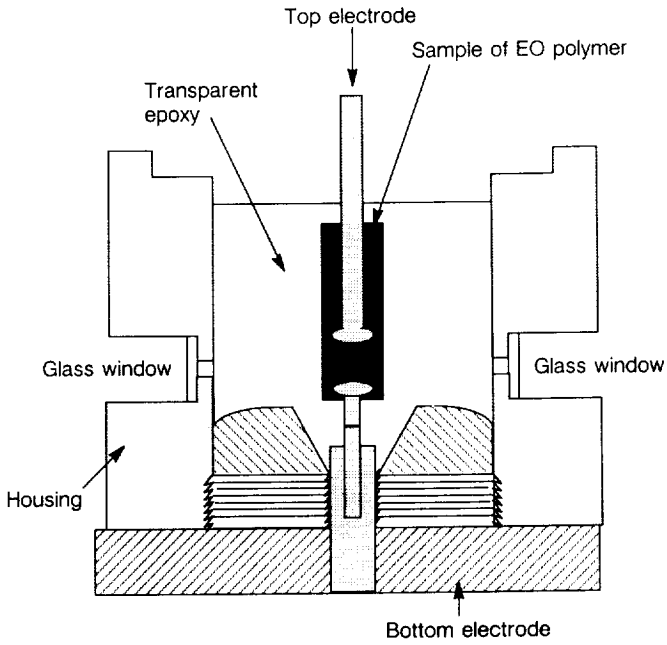


Figure 2-13 The second sensor mold

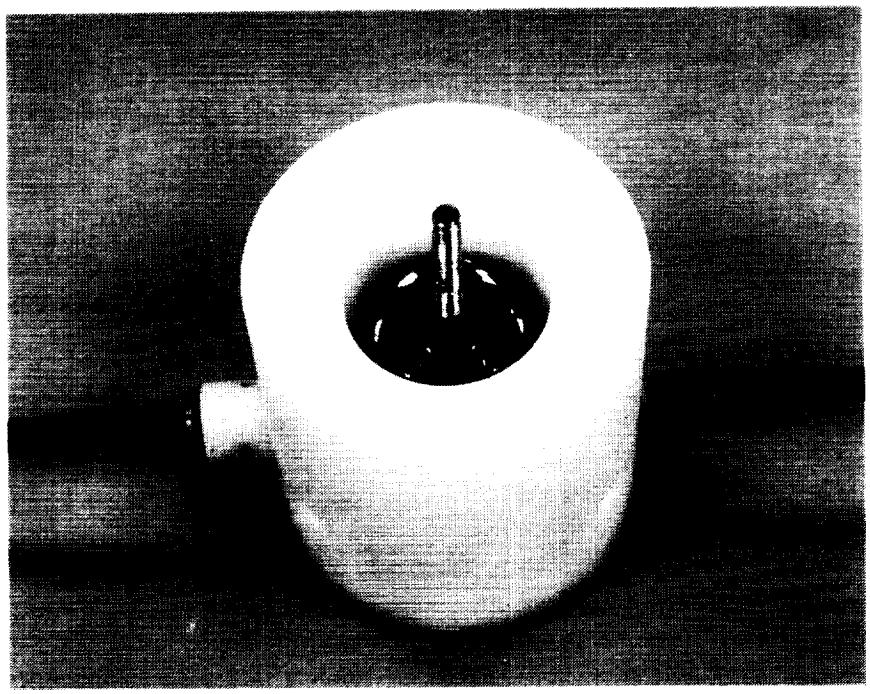


Figure 2-14 The voltage sensor, assembled.

2.3.2 Sensor design parameters

In order to optimize the sensor design, some design parameters deserve greater consideration. In the following discussion, it is assumed that the materials are those described earlier, so the material parameters are fixed.

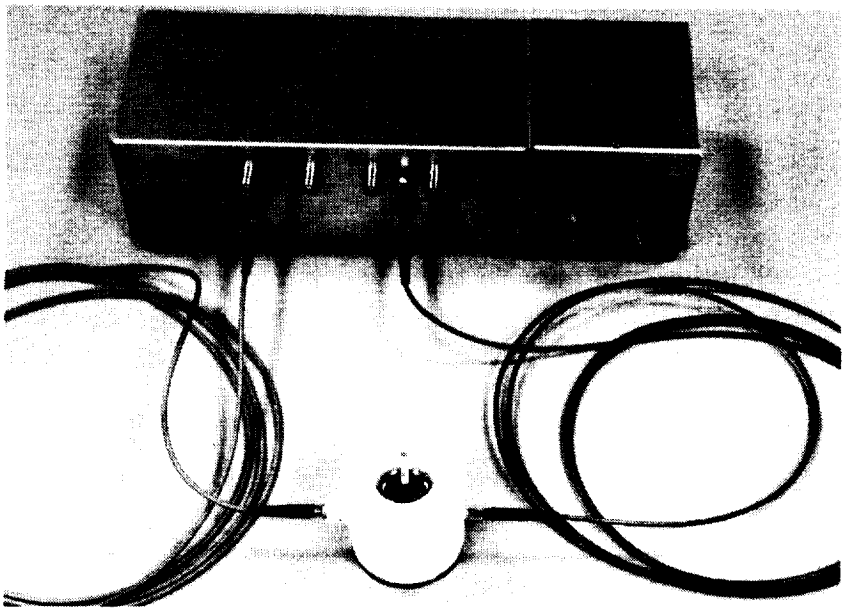


Figure 2-15 The voltage sensor with its driver electronics.

2.3.2.1 The electrode gap distance

To maximize the electro-optic effect in the polymer, the sample should be poled as close as possible to the field strength where dielectric breakdown occurs in the material. If the poling voltage is adjusted for maximum poling field, the electro-optic coefficients will be determined by the dielectric breakdown strength of the polymer and the dye concentration, and as such will only depend on the materials. Therefore, the electro-optic coefficients are independent of the electrode gap distance. The electrode gap distance thus determines only the breakdown voltage. The maximum voltage at which the sensor can be safely operated is a function of this voltage, and therefore, the maximum voltage for the sensor is determined by the electrode gap distance.

Larger maximum voltage requires larger gap distance.

Assuming that we have a constant noise level from a given optical transmission system and a given light source, the electrode gap will determine the resolution, which is defined as the voltage across the sensor for which the signal level is equal to the noise level (SNR = 1).

Finer resolution requires smaller gap distance.

2.3.2.2 The optical transmission system

Once the electrode distance has been determined, the main purpose of the optical design is to maximize the absolute sensor response, defined as

$$\text{ASR} = \frac{P_{ac}}{P_0 V} \tag{2.1}$$

where P_{ac} is the RMS value of the AC part of the transmitted optical signal, P_0 is the total optical power injected into the system and V is the voltage across the sensor.

It should be emphasized that it is the absolute sensor response and not the relative sensor response which should be maximized. The relative sensor response is defined as

$$\text{RSR} = \frac{m}{P_0 V} = \frac{P_{ac}}{P_{dc} P_0 V} = \frac{\text{ASR}}{P_{dc}} \tag{2.2}$$

where P_{dc} is the DC part of the transmitted optical signal, and m is the modulation index.

It is important to maximize ASR and not RSR because it is possible to have a very high modulation index and still a poor signal because little optical power is transmitted. If the signal to noise ratio is low, the resolution will be poor, no matter how high the modulation index is.

2.3.2.3 How to maximize the Absolute Sensor Response

In order to determine a mathematical expression for ASR, the following assumptions have been made.

1. The whole sample of electro-optic polymer is homogeneously poled and all of it is exposed to a homogeneous modulation field.
2. The only loss mechanisms dependent on the properties of the polymer sample are the absorption in the electro-optic dye and the coupling loss between the two GRIN rod lenses, the latter of which is a function of the distance between the lenses.

Transmitted optical power in a homogeneous non-absorbing medium between two GRIN rod lenses is determined by the coupling loss between them. It depends on the types of GRIN rod lenses, optical fibers and LED, and on the distance L_n between the lenses where the subscript n indicates the index of refraction of the material between the lenses.

Because it is difficult to deduce an analytical expression for the transmitted optical power, this quantity has instead been measured for different combinations of optical fibers and GRIN rod lenses but, with the same LED as light source.

The measurements were made in air, so the distance between the GRIN rod lenses was measured in air, and is represented by L_1, because the index of refraction for air is 1.00. The equivalent distance in a material with index of refraction n can be determined by the expression

$$L_n = nL_1 \tag{2.3}$$

Transmitted optical power $G(L_n)$ is shown in Figure 2-16 for different combinations of GRIN rod lenses and optical fibers. It is assumed that the average index of refraction is not very different for that of the transparent epoxy. Because of this assumption L_n has been calculated for the index of refraction of the transparent epoxy which is $n = 1.564$. The distance L_n between the GRIN rod lenses is the sum of the dimension l of the electro-optic polymer in the direction of the optical path, and a dead space t, occupied by other materials.

$$L_n = l + t \tag{2.4}$$

In the calculations, the transmitted optical power will be represented by $G(l+t)$.

When it is assumed that there is no reflection loss present, the transmission T through a sample of electro-optic polymer is represented by

$$T = e^{-\alpha l} \tag{2.5}$$

where α is the absorption coefficient.

The absorption A of the sample is defined as $-\log(T)$, and is related to the concentration c of the electro-optic dye in the sample, the dimension l in the direction of the optical beam and the extinction coefficient ε by means of the expression

$$A = -\log(T) = \varepsilon c l \tag{2.6}$$

From equations (2.5) and (2.6) can be derived the relation

$$\alpha = \ln(10)\varepsilon c = \frac{A}{l}\ln(10). \tag{2.7}$$

The transmitted optical power P_{dc} through the sensor can be expressed by the following relation:

$$P_{dc} = g e^{-\alpha l} G(l + t) \tag{2.8}$$

where g is an attenuation coefficient which represents the total loss originating from all loss mechanisms other than the absorption in the electro-optic polymer and the coupling loss between the GRIN rod lenses.

The relation

$$P_{ac} = m P_{dc} = m g e^{-\alpha l} G(l + t)$$

is substituted in equation (2.1), leading to the relation

$$\mathrm{ASR} = \frac{m g e^{-\alpha l} G(l + t)}{P_0 V} \tag{2.9}$$

Now α is substituted in expression (2.9) by expression (2.7). The relation $m = \Gamma$ from (A.3.7) is being used, where Γ is given by (A.6.6).

These substitutions change (2.9) to

$$\mathrm{ASR} = \frac{\pi n_0^3 (r_{33} - r_{13})}{P_0 \lambda d} l g e^{-\ln(10)\varepsilon c l} G(l + t) \tag{2.10}$$

We now use relation (A.5.19) between the electro-optic coefficients and the concentration of the electro-optic dye. By using this relation (2.10) can be brought into the form

$$\mathrm{ASR}(l, c) = p l c e^{-\ln(10)\varepsilon c l} G(l + t) \tag{2.11}$$

where p and t are assumed to be constants, and $\mathrm{ASR}(l, c)$ indicates that ASR is a function of the variables l and c.

In order to maximize ASR(l, c) we will divide it in the two parts $G(l + t)$ and $f(l, c)$ where the latter is defined by

$$f(l, c) = plce^{-\ln(10)\varepsilon cl} \tag{2.12}$$

It will now be demonstrated that these two functions are related in such a way that they can be maximized independently to provide the maximum for ASR(l, c).

It can be seen from Figure 2-16 that $G(L_n)$ is a decreasing function of L_n. As $L_n = l + t$, where t is a constant, it is evident that l should be as small as possible in order to maximize $G(l_n)$.

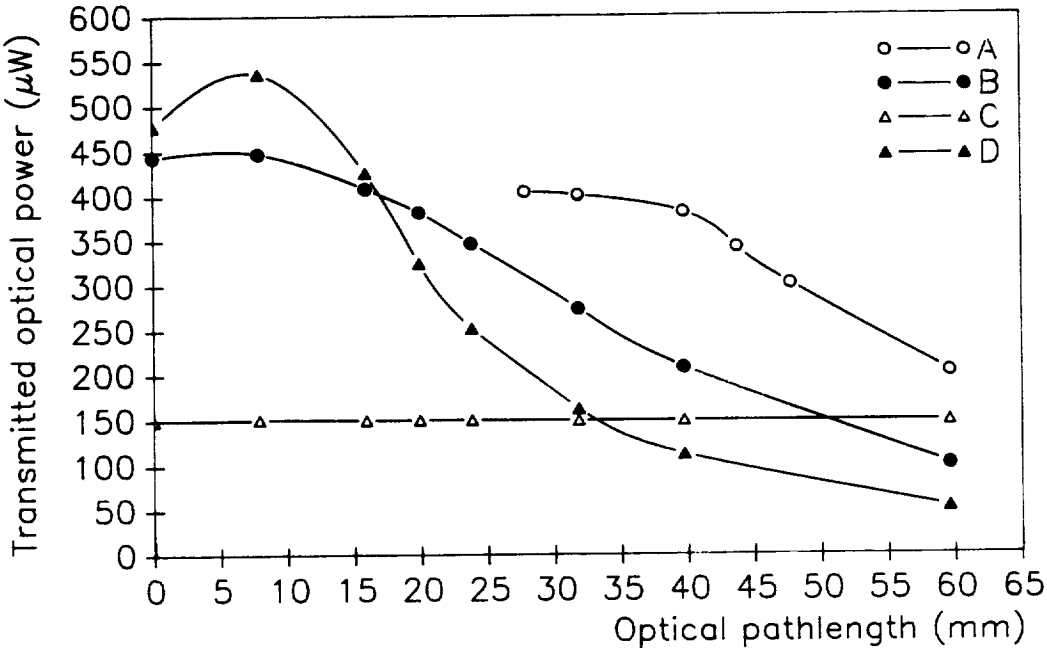


Figure 2-16 Transmitted optical power $(G(L_n))$ as a function of path length for different combinations of optical fibers and GRIN rod lenses. The data have been recorded with air as optical medium, but the path length has been corrected, so the data fits the refractive index $n = 1.564$.

The function $f(l, c)$ can be maximized by performing partial differentiation with respect to l and c, and searching for maxima among the value sets (l, c) for which both derivatives are zero.

Partial differentiation of $f(l, c)$ with respect to l and with respect to c leads to the two derivatives

$$\frac{\partial}{\partial l} f(l, c) = pce^{-\ln(10)\varepsilon cl}[1 - \ln(10)\varepsilon cl] \tag{2.13}$$

$$\frac{\partial}{\partial c} f(l, c) = ple^{-\ln(10)\varepsilon cl}[1 - \ln(10)\varepsilon cl] \tag{2.14}$$

Both derivatives vanish when l and c are related through the equation

$$\ln(10)\varepsilon cl = \alpha l = 1 \tag{2.15}$$

and it can be shown that when this expression is valid, $f(l, c)$ will obtain its maximum. It should be emphasized that $f(l, c)$ can obtain the same maximum value for an infinity of combinations of (l, c) as long as equation (2.15) is valid.

In the case where GRIN rod lenses are used, the dimension l of the polymer should be kept as small as possible in order to maximize the DC optical transmission $G(l + t)$. However, the maximum amount of dye which can be dissolved in the resins of the epoxy sets a minimum limit to the dimension l, when one wants relation (2.15) to be observed.

The first step in the process of maximizing the Absolute Sensor Response (ASR) is to select a dye concentration as high as possible and adjust the dimension l of the electro-optic polymer, so equation (2.15) is observed.

The next step is to select a combination of GRIN rod lenses and optical fibers for which $G(l + t)$ is as high as possible. Of course, the dead space t should be as short as possible, taking the necessary electrical insulation properties into account.

When the mold was first designed, the relationship (2.15) had not been discovered. However, it seems that the combination of dye concentration and physical dimension is close to ideal for the sample used in the sensor. The sample used a concentration of 5% (by weight) which is the maximum concentration of DR1 that can be dissolved in the resins of the epoxy. The absorbance of the 5%, 10-mm sample was measured at 820 nm. The absorbance measured was

$$A = 0.355 \tag{2.16}$$

Through the relationship (2.6) the extinction coefficient ε was calculated. The optimal dimension for a sample of 5% concentration can calculated by using (2.15).

The optimal dimension is

$$l = 12.2\,\text{mm}$$

It should be noted that the above does not differ significantly from the sample dimension of 10 mm. Unfortunately, the dimensions of the electrode housing had not been optimized for this dimension of the electro-optic polymer. With the 10-mm dimension of the electro-optic polymer, the optical path length could be made much shorter by cutting down on the thickness of the transparent epoxy around the sample, and still maintain enough wall thickness for electrical insulation. With a shorter pathlength a higher value of $G(l + t)$ would have been obtained.

2.3.3 Data on fiber optic HV sensor

In this section, data and test results will be provided on the demonstration voltage sensor built at JPL. The sensor performance will be evaluated by comparing the test results to theoretical values.

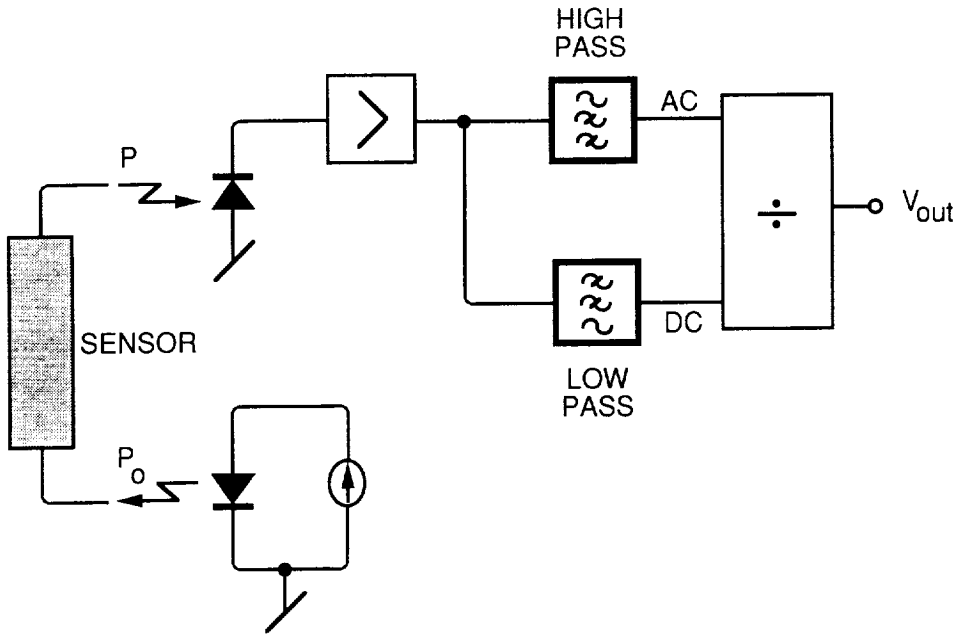


Figure 2-17 Signal detection circuit

2.3.3.1 Components in the sensor system

In the sensor head, the following components have been used:

Electro-optic polymer : Fabricated at JPL as described in this report, 5% DR1 in EPO-TEK 301-2, electrode gap 3 mm, length 10 mm, poled at 54 kV.

1/4 wave plate : Meadowlark Optics (CO), Achromatic retarder, center wavelength 850nm, diameter .5 in. The 1/4 wave plate consists of several layers of birefringent polymer material, cemented between optical windows.

Polarizers : Corning, 8612 Polarcor 800 High Contrast, nominal wavelength 800 nm, diameter 6 mm. (Polarcor is a mixed alkali borosilicate glass containing submicroscopic silver particles aligned along a common axis. Resonant absorption by the silver conduction electrons provides preferential absorption of the polarization component aligned with the long axis of the silver particles.)

GRIN rod lenses : Nippon Sheet Glass Company (Japan), [Vendor: Melles Griot (CA)],SELFOC, Melles Griot product number 06 LGS 212, pitch 0.25, N.A.0.37, diameter 2.0 mm.

For transmission of optical power the following were used:

Transmitter fiber : Siecor, core 200 μm.

Receiver fiber : 3M EOTec, TECS FT(400)LMT, core 400 μm, N.A. 0.37, max. attenuation at 850 nm is 6 dB/km.

The light source and the detector are a Light Emitting Diode and a PIN photodiode.

LED: Hewlett Packard, HFBR 1404, center wavelength 820 nm.

PIN diode: Hewlett Packard, HFBR 2208.

The signal detection circuit is shown in Figure 2-17.

2.3.3.2 Evaluation of the sensor response

The LED was driven by a stable current. First the optical power output P_0 of the transmission fiber was measured. Next the transmitter fiber and the receiver fiber were connected to the sensor, and the optical power P_{dc} at the far end of the receiver fiber was measured.

With a forward current of 70 mA through the LED the following values were measured.

$$P_0 = 588\mu\text{W} \tag{2.17}$$

$$P_{dc} = 22.0\mu\text{W} \tag{2.18}$$

With these constant optical power levels, the modulation index m was measured when the sensor was subjected to a voltage of 7.96 kV (RMS). The measured value was

$$m_{(7.96\text{ kV})} = 2.05 \times 10^{-3} \tag{2.19}$$

These values yield a measured absolute sensor response ASR1 equal to

$$\text{ASR1} = \frac{m \cdot P_{dc}}{P_0 \cdot V} = 9.64 \times 10^{-9} V^{-1} \tag{2.20}$$

Assuming that the only loss present is due to the absorption in the dye, the polarization loss and the coupling loss between the GRIN rod lenses, the absolute sensor response for this calculation can be calculated by using expression (2.7) in equation (2.9) yielding

$$\text{ASR2} = \frac{m \cdot \frac{1}{4} \cdot e^{-Aln(10)}}{P_0 \cdot V} G(L_n) \tag{2.21}$$

The length between the GRIN rod lenses is approximately 44 mm. It has been measured that with the combination of GRIN rod lenses and fiber used, $G_{(44mm)} = 342\mu\text{W}$. When using these numbers one obtains

$$\text{ASR2} = 1.65 \times 10^{-8} V^{-1} \tag{2.22}$$

It would be interesting to compare ASR1 and ASR2 to the maximum obtainable value ASR3 for the absolute sensor response, providing the materials are not changed, the light source is the same, and GRIN rod lenses are still being used in the optical transmission system.

It is assumed that the minimum distance obtainable between the GRIN rod lenses is 17 mm where the sample allows for 12.2 mm, and the rest will provide for electrical insulation. In this case it would be beneficial to use 400 μm fiber for the transmitter and the receiver fiber. In this case it was found $G_{(17mm)} = 420\mu\text{W}$. In the ideal case expression (2.15) is valid, and the expression for ASR3 is

$$ASR3 = \frac{1.22 \cdot m \cdot \frac{1}{4} \cdot e^{-1}}{P_0 \cdot V} \cdot G(L_n) \tag{2.23}$$

where the factor 1.22 is the ratio of the ideal pathlength to the pathlength in the physical sensor. When substituting the numbers for the parameters one gets

$$ASR3 = 2.06 \cdot 10^{-8} V^{-1}$$

This number is the maximum obtainable number for the analyzed combination of electro-optic material, light source, electrode gap and type of fiber optic transmission system. If one wants to increase this number, the following actions can be taken:

Electro-optic dye: Higher $\mu\beta$ product, less absorption, mixable with the polymer in higher concentrations.

Polymer: Higher dielectric strength, so higher strength of the electric poling field can be applied.

Electrode gap: Shorter electrode gap distance.

Light source: Higher optical output into transmitter fiber.

Optical transmission system: Capable of transmitting more optical power, less loss in system.

2.3.3.3 Resolution of the sensor system

The electronic detection system which was built proved to be far more sensitive than the system used on the optical bench. The noise level in the electronic system equals about 3.0 V across the sensor which consequently is the resolution of the system. The angular resolution can be calculated from

$$m(7.96kV) \cdot \frac{3V}{7.96kV}$$

and corresponds to an angular resolution of 0.8μ Radians. This number was far better than expected. With a resolution of this magnitude, it should be possible to use the sensor at much smaller voltage levels than the 20 kV it was designed for. If the electrode gap is made smaller, the resolution could be further improved.

2.3.3.4 Linearity and maximum voltage for the sensor system

The sensor system is expected to show excellent linearity from the minimum detectable voltage to very high voltage levels. The linear range of the sensor depends on how big a difference $(\Gamma - sin\Gamma)$ one will tolerate. Assuming that one would tolerate a deviation of 0.3% , one would define the linear range as the range from the minimum detectable voltage to the voltage where $\Gamma(V) - sin\Gamma(V) = 0.003$. This would equal an angle of approximately 260 m radians which would yield a theoretical maximum voltage of approximately 1 MV. As a result of this, it can be concluded that the maximum voltage of the sensor is not limited by nonlinearity, but by dielectric breakdown.

The sample of electro-optic polymer was poled at 54 kV (DC).This is believed to be close to the breakdown voltage of the sensor. Assuming that 54 kV is also the standard basic impulse the sensor can withstand, this would, according to the Basic Impulse Insulation Levels, only qualify it for a Reference class of 2.5 kV in the case of distribution class equipment. However, if the Basic Impulse Level the sensor can withstand is higher than the DC breakdown strength, it may qualify for a higher class.

For demonstration purposes, it should be safe to operate the sensor at voltage levels of 10–20 kV with a large current limiting resistor in series. The prototype sensor is not expected to be damaged at voltage levels below peak values of 54 kV, equaling 38 kV (RMS).

2.4. Proposals for Future Work

The purpose of this work has been to demonstrate that it is possible to fabricate electro-optic polymer in bulk form and use it for voltage measurements at high voltage levels. However, more work is required before the technology can be used for practical applications in commercial high voltage sensors or large aperture external laser modulators.

2.4.1 Materials test

In order to evaluate the material under development, a number of parameters should be studied.

2.4.1.1 Aging effects

Is there any degradation in the electro-optic effect with time? It has been reported for thermopoled samples of DR1 in PMMA that the electro-optic effect has decayed considerably in just a few days because of thermal relaxation in the polymer.

As the combination of DR1 in epoxy poled in the combined cure- and thermopoling process is new, no data have been recorded for this material. Epoxy is known to be a much more stable material than PMMA, and has a higher glass transition temperature, but a decay in the electro-optic effect with time may still be a possibility.

A sample of the poled electro-optic polymer has been tested with a time interval of close to one month and no decay in the electro-optic effect was detected. However, the sample had been removed from the test set up, and the adjustment of the position of the optical components could have changed as a result of this. In the fiber optic sensor, the optical components can be mounted in fixed positions, so misalignment can be eliminated.

After the fiber optic sensor has been completed, a long duration test program should be initiated to study the aging effect of the electro-optic polymer.

2.4.1.2 Temperature sensitivity

The electro-optic polymer can be expected to be temperature sensitive, because of the following mechanisms:

The electro-optic coefficients are temperature dependent.

The intrinsic birefringence, which is known to be present, is dependent on the linear dimensions of the material, which will change with temperature. The linear dimensions are included in the expression for the phase retardation, giving rise to another temperature dependent factor. These mechanisms account for temperature sensitivity in the electro-optic polymer. The optical polarization components are also temperature sensitive. It is hard to estimate the overall temperature dependency of all these effects. Some of them might cancel out, some of them will add up. With the sensor, it is easy to test what the temperature dependency is for the whole system. If the temperature dependencies of the other optical polarization components are known, the temperature dependency of the electro-optic polymer can be estimated from a temperature cycling experiment.

2.4.1.3 Linearity

According to theory, sensor response is expected to deviate very little from a linear approximation over the whole dynamic range. It should be verified that the system behaves according to theory.

2.4.2 Development of materials

The development of better materials is essential for the future of this technology. In order to make better sensors, especially for lower voltage applications, higher electro-optic modulation would be necessary. In order to obtain this materials with the following properties would be beneficial :

Electro-optic dyes

High $\mu\beta$ product

Good solubility in the resins of the polymer

Low absorption at wavelengths available in LEDs

Polymers

High dielectric breakdown strength

Low absorption at selected wavelengths, particularly in the visible spectrum

Comment: If it were possible to find an electro-optic dye transparent in the visible spectrum, say at 650 nm, it would be possible to use extremely low cost optical polarization components and LEDs. This could make production of very low cost sensors possible.

Other important properties for the electro-optic material, unrelated to electro-optic response, but related to the lifetime of the material are:

High glass transition temperature (good thermal stability)

Low thermal expansion

Little shrinking in the curing process

2.4.3 Development of fabrication process for EO polymer

It has been demonstrated that it is possible to fabricate electro-optic polymer in bulk form and use it in a voltage sensor at high voltage levels. However, in order for the fabrication process to be feasible, further development work should be done. The material could be fabricated in sheets, sandwiched between electrodes made of a thin metal foil (in the same way dry wall is made of gypsum covered with paper). The thickness of the sheet of electro-optic polymer would depend on the voltage level at which the material was going to be used. When the sheets had been made, individual pieces of polymer for the sensors could easily be cut.

In order to fabricate the electro-optic polymer in sheet form, a special poling apparatus should be developed. In the development phase the fabrication method could easily be tested with sheets of relatively small dimension. As vacuum is necessary to get the dissolved air out of the resins, to avoid dielectric breakdown in the cured polymers, the poling apparatus should be set up in a vacuum oven. During the poling process, the oven should be heated up to the desired temperature. Flooding the oven with an electrically insulating gas, at a pressure slightly higher than atmospheric pressure, could prevent dielectric breakdown in air. The correct poling apparatus design, combined with the use of electrically insulating gases, would probably eliminate the need for an oil bath in the poling process. During the poling process, temperature, voltage and current through the material should be controlled. An algorithm to control these parameters could be devised to put the whole process under computer control.

2.4.4 Development of sensors

Small pieces can easily be cut from a sheet of electro-optic polymer and mounted in a prefabricated sensor housing. The sensor housing can be made in such a way that the electro-optic polymer is automatically mounted in the correct position. The electro-optic polymer is then potted in a transparent polymer which is simply poured into the sensor housing. Because of the match in index of refraction between the electro-optic polymer and the transparent polymer, only very little polishing of the optical surfaces is necessary. The optical polarization components could be mounted in the sensor housing and be potted into the transparent polymer in the same manner. Even the tips of the optical fibers could be embedded in the transparent polymer. Several types of polarizers and 1/4 wave plates are currently on the market. The price and performance of the different types should be thoroughly studied, especially the temperature dependency of the components. As the electro-optic epoxy is intrinsically birefringent, it might even be possible to control the intrinsic birefringence, to provide for the 1/4 wave of phase retardation. The optical transmission system in the demonstration voltage sensor is quite sensitive to even a small amount of misalignment. This problem could possibly be solved by redesign. In addition, sensor installation on the power line should be studied.

One possibility would be to build the sensor into a hollow insulator and pot it into an insulating material. This concept has the disadvantage that one terminal on the sensor has to be directly wired to ground. A second possibility would be to place the sensor between the high voltage inner conductor and a conducting cylinder at a floating potential surrounding the inner conductor. This concept has the disadvantage that the potential of the surrounding cylinder is dependent on the ambient electric field. If the ambient electric field changes (as a truck goes by, or during rain) the output of the sensor will change. However, if the distance from the surrounding cylinder to what perturbs the ambient field is large compared to the distance between the inner conductor and the surrounding cylinder, the change in output will be small, and as a result of this, the sensor will still be fairly accurate.

2.5. Conclusions

The theoretical background for fabrication of electro-optic polymer in bulk form and the use of this type of material in a voltage sensor have been discussed. It has been experimentaly demonstrated for the first time that it is possible to fabricate electro-optic polymer in bulk form. Though the electro-optic coefficients obtained have been rather small, they have been large enough to utilize the material in a fiber optic voltage sensor operating over a range from a few volts to 35000 volts, with excellent linearity. As such, the concept has been proved that it is possible to fabricate electro-optic polymer in bulk form and use it in a fiber optic high voltage sensor. However, the long term stability and the thermal stability of the electro-optic polymer (important material properties) remain to be studied. Over the duration of this work, no weakening of the electro-optic effect was observed. Future chemical development of the materials can be expected to lead to higher dielectric breakdown strength, higher sensitivity and increased stability. It is expected that the fabrication process described in this report can be expanded to a larger scale production of electro-optic polymer. Electro-optic polymer could become the key material in future fiber optic voltage sensors.

Bibliography

Hecht, E. and Zajac, A. (1987). Optics, 2nd edition, World student series edition, Reading, MA, Addison-Wesley publishing company.

Yariv, A. and Yeh, P., (1984). Optical waves in crystals: propagation and control of laser radiation, New York, NY, Wiley & Sons.

Singer, K. D., Kuzyk, M. G. and Sohn, J. E., (1987). Second-order nonlinear-optical processes in orientationally ordered materials: relationship between molecular and macroscopic properties. J. Opt. Soc. Am. B/Vol.4, No. 6, June 1987, pp. 968-975.

Sohn, J. E., Singer, K. D., Kuzxy, M. G., Holland, W. R., Katz, H. E., Dirk, C. W., Schilling, M. L. and Comizzoli, R. B., Orientationally ordered nonlinear optical polymer films, AT&T Engineering research center, P. O. Box 900, Princeton, NJ, 08540, USA, AT&T Bell Laboratories, Murray Hill, NJ, 07974, USA.

SECTION 3.

POWER SYSTEM STATIC SECURITY ASSESSMENT USING THE KOHONEN NEURAL NETWORK CLASSIFIER

3.1 Introduction

An important task in power system operation is to decide whether the system is currently operating safely, critically or unsafely. If the system is safe, it is further of interest to know whether following an unforeseen outage of one or several lines, transformers or generators on the system will still remain in a safe operating state. Important criteria for security assessment are violation of constraints for voltages, power flows, islanding etc. Contingency analysis for the case of one or several presumed outages of lines or generators ($N-1$ or $N-M$ contingencies respectively) tests the constraint violation for load flows.

There exist powerful algorithms like fast decoupled load flow calculations which treat ($N-M$) contingencies completely (Stott, 1974). However, the systematic computation of all failures required by a $N-M$ security analysis (ie the simultaneous failure of M out of N elements) leads to an exponential increase of the computation time. For large networks, these multiple failures in the electrical network cannot be treated in real time if conventional computers are used to detect critical situations by simulations. Other methods like contingency ranking give fast results but do not detect all critical states (Ejebe, 1979).

Statistical pattern recognition techniques for security analysis were proposed to take human experience into account by Pang (1974). The processing power available at the time when these methods were developed was insufficient for their use in an industrial environment. More recently the application of associative memories for static security assessment of power system overloads has been studied at EPRI (1982) and by Oh (1986).

Today other Artificial Intelligence methods are studied for this purpose. Two main approaches to solve problems in power system analysis can be distinguished. Firstly, expert system techniques that explore human knowledge formulated explicitly as rules were investigated for static security analysis (Sobajic, 1987).

Secondly, artificial neural network techniques for security analysis have been proposed for the estimation of critical clearance time for the transient stability problem (Sobajic, 1988). For static security assessment, a multi-layered network and the back-propagation algorithm were applied to a 4 machine–8 bus–14 line system by Aggoune (1989). Classification of contingencies of a 4 machine–17 bus–35 line network has also been achieved with the Hopfield network (Chow, 1990).

However, approaches which simulate backpropagation networks are extremely time-consuming and necessitate a huge amount of training data. They do not take into account the properties of the system space but try to conceive the problem as a functional interpolation task. Due to the high dimensionality of the function being interpolated, accuracy is highly dependent on the number of training points. Further, each contingency has to be viewed as a different function. Therefore the combinatorial aspect of this task is still preserved. This argument is also valid for the Hopfield net approach. To our knowledge, none of the proposed models has actually been implemented in hardware.

The objective of the work reported here is to study Kohonen's self-organizing feature map for the classification of security states of power systems with respect to contingency analysis. The network determines similarities of system states in an unsupervised manner and successfully identifies line loading patterns for unknown as well as trained system states.

In contrast to the supervised learning of the backpropagation algorithm, unsupervised learning does not require the target class to be specified. To our knowledge the study of unsupervised neural nets for load forecasting is the only application of this technique to power system problems (Dillon, 1975).

Therefore in the next section a short introduction to Kohonen's self-organizing feature map is given. In section 3 the application of this classifier to power system security assessment is presented and simulation results are discussed.

3.2 Kohonen Self-Organizing Neural Network

3.2.1. Architecture of Kohonen network

In accordance with experimental results (see, for example, Hubel, 1977), it has been postulated that information in the brain is stored on a two-dimensional surface, and that related information occupies neighboring locations on that surface.

Mathematical models describing these self-organizing neural networks have been proposed initially by von der Malsburg (1973) and developed further by Kohonen (1982). Throughout this Section the terms Kohonen neural network, Kohonen classifier and self-organizing feature map are used as synonyms for this type of network.

There is no target output and no "teacher" for evaluating an error function in these self-organizing networks. The learning of the synaptic weights is unsupervised, which means that, upon presentation of new input vectors, the network determines these weights dynamically, such that input vectors which are closely related (in a sense to be precisely defined below) will excite neurons which are close together, or clustered.

The Kohonen network is an array of **M** processing elements (neurons). If these **M** neurons are arranged on a grid in a plane, the network is called two-dimensional, since this network maps high dimensional input vectors onto a two-dimensional surface. For a given network, the input vectors $\mathbf{x}$ have a fixed dimension **N**. The **N** components of the input vector $\mathbf{x}$ are connected to each neuron in the array. A synaptic weight w_{ji} is defined for a connection from the ith component of the input vector to the jth neuron. Therefore, an **N**-dimensional vector $\mathbf{w}_j$ of synaptic weights is associated with each neuron j. The neurons themselves are also connected in a predefined manner according to a given neighborhood relation.

Figure 3-1 shows a Kohonen network which maps 3-dimensional input vectors on a 2- dimensional map containing 16 neurons. A neuron in this square network has four direct neighbors, called neighbors of first order, except for the neurons along the borders of the grid. Two neurons i and j are neighbors of order k if i has a direct neighbor h who is a neighbor of j of order $k-1$.

The number of neurons does not depend on the dimension of the input vector nor on the size of the training set. However a small number of neurons can form only a small number of clusters. Therefore, depending on the application, the number of discriminated features might not be sufficient for a large training set.

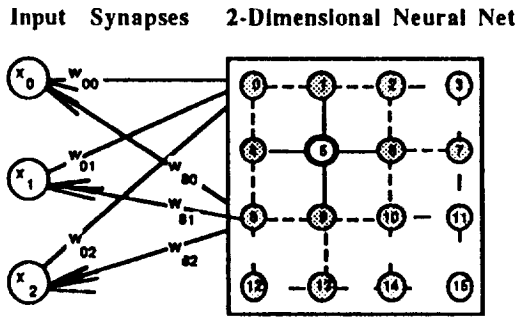


Figure 3-1. Kohonen network

3.2.2 The self-organization algorithm

Let M be the number of neurons with weight vectors

$$\mathbf{w}_i \in R^N, 1 \leq i \leq M$$

and

$$X = \{\mathbf{x} \in R^N \mid \mathbf{x}\text{trainingvector}\}$$

the training set. Define a neighborhood function $\alpha(t, i, c)$ such that $\alpha(t, i, c)$ is monotonically decreasing in training time t and in the order of neighborhood of neuron i and neuron c. Let $\varepsilon > 0$ be small.

Phase A: Unsupervised Learning

1: $t_0 := 0$: initialize all w_{ij} randomly
2: Choose input vector $\mathbf{x} \in X$ randomly
3: Determine the neuron c such that its weight vector $\mathbf{w}_c$ is the closest to the input vector

$$\| \mathbf{w}_c(t_n) - \mathbf{x} \| = \min \| \mathbf{w}_i(t_n) - \mathbf{x} \| \quad \text{for all } i$$

4: Increment the time $t_{n+1} = t_n + \Delta t$
5: Update the weight vector $\mathbf{w}_c$ of the selected neuron c and the weight vectors $\mathbf{w}_i$ of its neighbors

$$\mathbf{w}_i(t_{n+1}) = \begin{cases} \mathbf{w}_i(t_n) + \alpha(t_n, i, c)(\mathbf{x} - \mathbf{w}_i(t_n)) & \text{for } i \in \text{neighborhood of } c \\ \mathbf{w}_i(t_n) & \text{elsewhere} \end{cases}$$

6: if $\alpha(t_n, i, c) > \varepsilon$ goto 2 else STOP

At each presentation of an input vector, ie at each a step of the learning process, every neuron of the network calculates a scalar activation function which depends on the input vector and on its own weight vector $\mathbf{w}$.

This function is chosen to represent a distance $\| \ldots \|$ between the input vector and the weight vector of the neuron under consideration. Possible choices are the Euclidian distance or the scalar product. Furthermore, there is a feedback from the neighboring neurons to a given neuron, according to a proximity rule and the learning time represented by $\alpha(t, i, c)$. In the original work by Kohonen, the neighboring neurons have a positive feedback, whereas the more distant neurons have a negative feedback.

The principle is that the neuron with the weight vector closest to the input vector, and its neighbors up to a certain distance (a distance which decreases as the learning goes on), will be adapted when an input vector is presented.

Phase B: Classification

1. Get an input vector $\mathbf{x}$
2. Choose $\mathbf{w}_c$ such that

$$\| \mathbf{w}_c(t_{max}) - \mathbf{x} \| = \min \| \mathbf{w}_i(t_{max}) - \mathbf{x} \| \quad \text{for all } i$$

For each vector in the N-dimensional input space, whether presented during the training phase or not, there is one responding neuron in the two-dimensional array of the Kohonen network, which has been excited most since its weight vector is the closest to the input vector (in the sense of a specified distance). One neuron may respond to several input vectors. Input vectors are mapped onto the weight vectors.

It was shown by Kohonen (1989) that this mapping preserves the topological order of the input vectors which is not necessarily known *a priori*. In fact the weight vectors themselves are constructed as weighted sums of a certain number of input vectors. If the input vectors form a vector space (in the mathematical sense) the weight vectors are elements of the same space representing a set of vectors of the same region of the vector space.

The choice of the input vector, the number of neurons and the neighborhood function will be discussed in section 3 with respect to the application to security analysis.

The adaptation process can be modeled by a Markov process. Stability and convergency properties of the learning algorithm of the self-organizing feature map have been studied in Cottrell (1986) and Ritter (1988). Analog implementations of self-organizing feature maps are examined by Vittoz (1989). The implementation of a neuron performing the self-organization mapping with learning capabilities in CMOS technology is discussed by Hochet (1990).

3.2.3 Technical applications

Kohonen networks have successfully been applied in robotics to learn the inverse kinematic transformation of a robot with two degrees of freedom by Ritter (1987).

A promising application of self-organizing network is the representation in two dimensions of input vectors of larger dimension. That technique has been used for the design of integrated circuits by Tryba (1988).

3.3 Application of the Kohonen Classifier to Security Assessment

3.3.1 Representation of the operating space

In this section the problem of static security assessment is formulated as a classification task. The operating state is described by a vector consisting of the active and reactive power of the power transmission lines. This vector is an element of an N-dimensional complex vector space called operating space where N is the number of lines. The operating point of a single contingency is an element of an N-1 dimensional subspace. The general concept of security regions and other definitions of operating points have been studied by Wu (1982) and Fischl (1982).

The boundaries of the secure state space are given by the limits of the buses and lines according to different security criteria. This Section treats the case of line overload prevention where the limits are given by the maximal power supported by these lines: $\| P_{ik} + jQ_{ik} \| \leq S_{ik\,max}$, where $P_{ik} + jQ_{ik}$ is the complex power flow.

A safe operation point always lies inside the boundaries of the secure space. Critical points lie inside but near the boundaries and unsafe points are outside the boundaries. In case of the linear power system model the secure region of the operating space is convex. Figure 3-2 illustrates this concept.

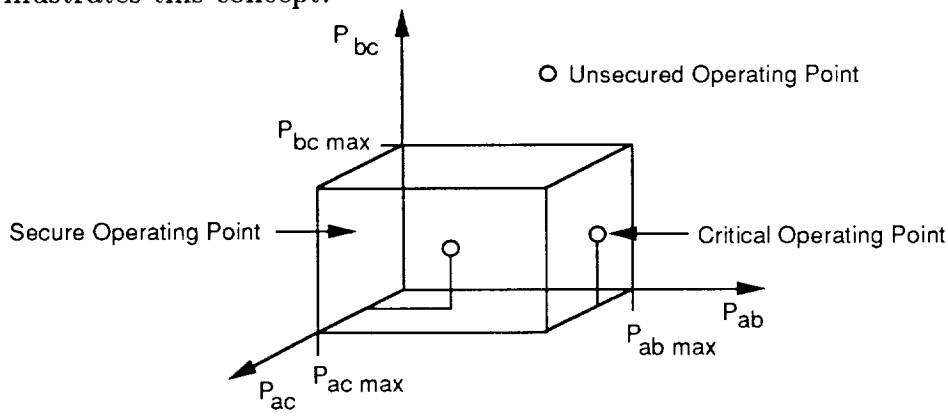


Figure 3-2. Operating space of a 3-bus 3-line linear power system model

3.3.2 Simulation of a 5-bus 7-line power system

The Kohonen network is used to classify line loading patterns resulting from single and double contingencies for a 5-bus 7-line power system defined by Stagg (1968) and represented in Figure 3-3. The line parameters and the generation and load at each bus are shown in Table 3-1.

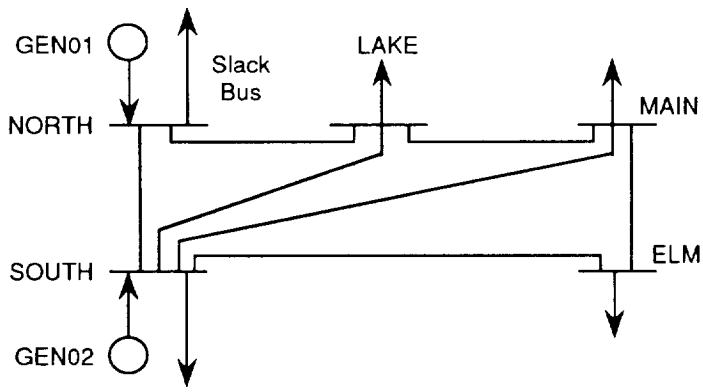


Figure 3-3. 5-bus 7-line system

For 7 lines the system state is defined by 7 complex, respectively 14 real, components; the 7 active and reactive line power flows. These input vectors were obtained by off-line load-flow simulations. One load and generation scenario was defined for the base case. The base case - 7 single line contingencies, 21 double line contingencies, 2 single generator contingencies, 1 double generator outage and 2 × 7 single-generator single-line outages - was simulated forming a set of 46 input vectors.

Table 3-1. Bus Data

Bus	Generation		Load	
Name	PG	QG	PL	QL
NORTH	(Slack	bus)		
SOUTH	40.	30.	20.	10.
LAKE	45.	15.		
MAIN	40.	5.		
ELM	60.	10.		

Reallocation of power on the generators was not performed, with the following exceptions:

- in the case of outage of one generator, the missing power is reallocated to the other generator.
- in some cases of double branch contingencies, where one or several loads cannot be supplied, the generator power is rescheduled.

3.3.3 Simulation data

A simulator for the Kohonen learning algorithm was developed in the language C on an Apollo DN 3500 work station. All the parameters can be defined and changed interactively: size of the input vector and of the Kohonen network, neighborhood function parameters and their decrease with the learning steps.

Vectors are drawn randomly from the set of 46 training vectors, and are presented as inputs to a two-dimensional Kohonen network with 49 neurons. The neurons are arranged on a 7×7 grid and numbered, row by row, from 0 to 48. The components of the weight vectors associated with each neuron have been randomly initialized between 0 and +1 MW and 0 and 1 MVAR respectively. Further, we define a *neighborhood order* function

$$k(t) := \lfloor k_0 \exp(\frac{-t}{\tau_k}) \rfloor$$

where $\lfloor x \rfloor$ denotes the largest integer less or equal to x. The *neighborhood set* $Nd(t, c)$ of neuron c is determined such that neuron $i \in Nd(t, c)$ if i is a neighbor of c of order d with $d \leq k(t)$.

Finally we define the *neighborhood* function

$$\alpha(t,i,c) := \begin{cases} \frac{\alpha_0 \exp(\frac{-t}{\tau_\alpha})}{1+d} & \text{for } i \in Nd(t,c) \\ 0 & \text{elsewhere} \end{cases}$$

The constant values have been chosen as

$$k_0 = 49, \tau_\kappa = 1000, \alpha_0 = 1 \text{ and } \tau_\alpha = 1000.$$

After 4000 steps this network is well organized, ie the weight vectors do not change significantly.

All parameters have been determined experimentally. Satisfying results have already been achieved with a 4×4 network and faster decreasing neighborhood functions. However it is evident that the number of clusters can not exceed the number of neurons. Further a slow decreasing neighborhood function will enhance the generalization capabilities of the network since for a large number of steps all neurons will be adapted to all input vectors. The number of neurons of the example network has not been chosen to equal the square of the number of lines. The interaction between neural net size, neighborhood relationship and power system size will be explored in future work.

3.3.4 Interpretation of the results

3.3.4.1 Classification of training vectors

Table 3-2 shows the classification results for 46 training vectors. For the table the following notations have been used: All buses are labeled by their first letter, all lines by the pair of buses which they connect. Therefore N denotes generator outage at bus North, S denotes generator outage at bus South, N-S denotes a single line outage of line North-South, N-S & L-M denotes a double line outage at line North-South and line Lake-Main and N-S & S is the double contingency of line North-South and generator outage at bus South. The "Base Case" denotes the case without any line or generator outage. S & N is the case where no generator is working and thus no load is supplied. Neurons containing empty cases do not classify any of the training vectors. That means their weight vector has never been chosen as the closest vector to any of the training vectors.

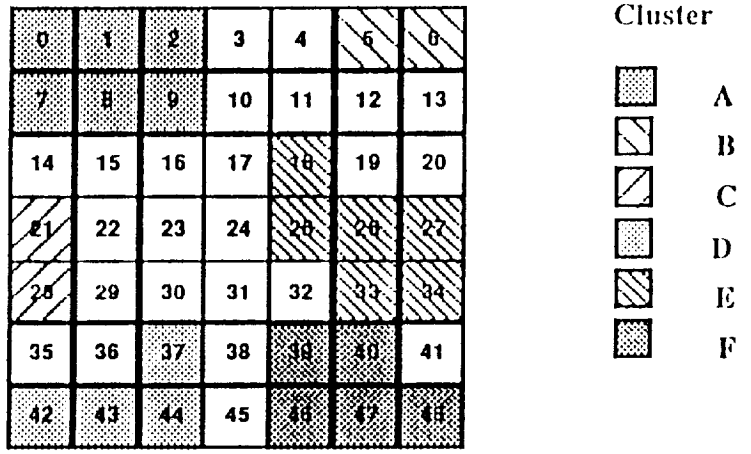
Table 3-3 illustrates the 6 different clusters present in Table 3-2. The numbers correspond to neurons 0 to 48.

Six different clusters can be distinguished. Four out of six are surrounded by "empty" neurons (like neuron 3) which do not classify any training vector. These clusters can be clearly distinguished from the neighboring clusters. The decision that the regions of cluster E and F form 2 different clusters and not 1 or even 3 has been made based on our interpretation of the classified cases. These 2 clusters do merge gradually into another.

Table 3-2. Classification of 46 cases by 49 neurons

Neuron 0 N-S & L-M	Neuron 1 N-S & S-L	Neuron 2 N-S & M-E	Neuron 3	Neuron 4	Neuron 5 S-E & M-E	Neuron 6 S & N N-L & N-S
Neuron 7 N-S & S	Neuron 8 N-S & S-M	Neuron 9 N-S N-S & N	Neuron 10	Neuron 11	Neuron 12	Neuron 13
Neuron 14	Neuron 15	Neuron 16	Neuron 17	Neuron 18 S-M & S-L	Neuron 19	Neuron 20
Neuron 21 N-S & S-E	Neuron 22	Neuron 23	Neuron 24	Neuron 25 S-M S-M & N S-M & M-E	Neuron 26 Base Case N M-E	Neuron 27 S-L S-L & N S-L & L-M S-L & M-E
Neuron 28 S-E & S-M	Neuron 29	Neuron 30	Neuron 31	Neuron 32	Neuron 33 S-M & L-M	Neuron 34 S-L & S L-M L-M & N L-M & M-E
Neuron 35	Neuron 36	Neuron 37 S-E & S	Neuron 38	Neuron 39 S-M & S	Neuron 40 L-M & S	Neuron 41
Neuron 42 S-E S-E & N	Neuron 43 S-E & L-M S-E & S-L	Neuron 44 S-E & N-L	Neuron 45	Neuron 46 S M-E & S	Neuron 47 N-L & S-M	Neuron 48 N-L N-L & N N-L & S N-L & S-L N-L & L-M N-L & M-E

Table 3-3. Cluster map for classification of 46 trained vectors



A Cluster A groups 7 out of 9 cases where line North-South is out of service. These cases correspond to a loading pattern where line North-Lake is always overloaded and line Main-Elm is weakly loaded (less than 30%). Therefore the first component of the weight vector corresponding to the active power of line North-Lake is significantly greater than the other 13 components. The components 3 and 4 of the input vector corresponding to the active and reactive power of line North-South are both zero since North-South does not carry any power. Therefore components 3 and 4 of the weight vector are relatively small. Finally components 13 and 14 corresponding to a weak load in Main-Elm are also small.

The remaining 2 cases of North-South outages have been classified by neuron 6 of cluster B and neuron 21 of cluster C respectively.

B Cluster B represents the set of double outages where the load cannot be supplied totally. These are the special cases where bus North or Elm are isolated from the rest of the network and the case of the double generator outage. However, these cases do not represent an overload situation and their loading pattern is quite different from those of cluster A.

C Cluster C corresponds to a loading pattern where as in cluster A line North-Lake is heavily loaded, even overloaded. But in contrast to cluster A line Main-Elm is loaded around 66%. Further, line South-Elm is out of service in both cases.

D Cluster D classifies cases where South-Elm is out of service and Main-Elm is loaded around 66%. In contrast to the cases of cluster C, North-Lake is normally loaded (up to 50%). North-South is either heavily loaded or overloaded (72% to 135%). South-Main is loaded around 50% for neurons 37 and 42, heavily loaded to overloaded for neurons 43 and 44. The decision that 37, 42, 43 and 44 form 1 cluster and not 3 depend on the importance of the loading pattern of line South-Main.

E Cluster E represents the non-critical cases where no overload is present and line Main-Elm is very weakly loaded (less than 10%).

One exception should be noted here: Contingency case South-Lake & generator South, classified by neuron 34 represents an overload of line North-South. For line South-Lake neurons 21, 27 and 43 handle the normal cases and neurons 34 and 48 the overloads. The fact that different neurons classify these cases shows that the behavior of line South-Lake is not an important feature of the network. This corresponds to its location in the network. Generally its outage does not provoke major changes.

For outages of line Lake-Main, cluster E also handles the non-critical cases whereas all critical cases provoking an overload of line North-South are classified by neurons 40 and 48 of cluster F.

F In general cluster F represents loading patterns of weak load of Main-Elm and overload of North-South. Thus 7 out of 9 outages of line North-Lake corresponding to these patterns are classified by cluster F. (The other 2 outage cases of North-Lake & North-South and North-Lake & South-Elm do not fall in this category!)

To interpret the organization of the weight vectors, we use the technique proposed in Tryba (1988).

Some of the neurons in the network are associated with several input vectors, and some to none of the input vectors. However, the weight vectors for those neurons appear to represent feasible states of the power system, which can be understood as generalization from the learned vectors. The weight vectors themselves represent the weighted sum of a certain number of input vectors which form a cluster. Therefore, with our choice of variables for the components of the input vectors, each component of a weight vector represents either an active or a reactive line power flow. For example, the first component of every weight vector corresponds to the active power in the branch North-Lake. Table 3-4 shows the data for the weight vector of neuron 9 classifying the single outage North-South.

The Euclidian distance between weight vector and input vector (= 12.718) is the minimal distance between this input vector and any weight vector of the network.

If we assume an active power limit in each of these lines, the comparison of this limit to the corresponding weight vector component of neuron i indicates whether neuron i classifies system states that are likely to violate this limit or not. Figure 3-4 illustrates this concept for the active power of line North-Lake. The values of the first component of all 49 weight vectors are presented on a map, at the location of the corresponding neuron defined in Table 3-2. In Figure 3-4 neurons occupy the edges of the squares.

Table 3-4. Weight vector $(\mathbf{w})_{2j+i} (1 \leq j \leq 7, i = 0, 1)$ of neuron 9 and line power flow for outage North-South

w 9 2j	w 9 2j+1	LINE	p [MW]	q [MVAR]	loading [%]
144.758	2.107	NORTH-LAKE	144.73	-3.66	131.79
3.262	-1.128	NORTH-SOUTH	0.00	0.00	0.00
42.248	16.632	SOUTH-ELM	39.36	17.47	39.20
-1.170	23.280	SOUTH-MAIN	-3.97	23.80	21.96
-19.033	34.136	SOUTH-LAKE	-15.38	29.91	30.62
61.807	-27.470	LAKE-MAIN	66.87	-32.62	71.07
19.397	-8.063	MAIN-ELM	21.89	-11.20	23.41

The features of these clusters are directly represented by the values of the weight vectors. For example large values of the active power of line North-Lake correspond to large values of the corresponding first component of the weight vector. Figure 3-4 illustrates how these features can be detected.

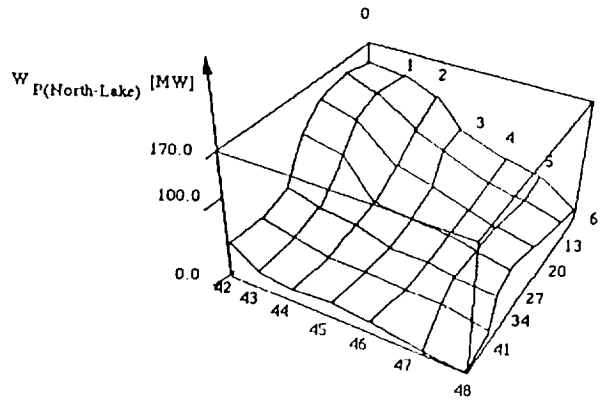


Figure 3.4. First component of the 49 weight vectors corresponding to the active power flow of line North-Lake

The limit for line North-Lake is at 100 MW. Therefore neurons 0, 1, 2, 3, 7, 8, 9, 10, 14, 15, 16, 21, 22 and 23 represent overload values whereas the values at neurons 47 and 48 corresponding to North-Lake outages are small. Also operating states with a heavily but not overloaded line North-Lake may be classified by a neuron whose weight vector indicates an overload situation. The Euclidian distance between this vector and the weight vector can be used as a measure of how likely this case will occur. Note that the operator in a utility faces a similar dilemma. There is no guarantee that a line loaded 99% will definitely not fail and that a line loaded 102% will definitely fail.

3.3.4.2 Classification of unknown vectors

In order to generate vectors which have not been trained but which are a "reasonable" distance from the 46 trained vectors, load and generation of the base case were uniformly changed to 90%, 95%, 105% and 110%. The 4×46 vectors obtained were presented to the Kohonen network trained on the 100% case. The claim of a reasonable distance is justified by the practical consideration that load and generation for a network usually vary around a scheduled case.

Instead of presenting in detail all 184 cases the main results will be presented.

1 All 6 clusters are still present and even more pronounced. For example, neuron 28 now classifies all 4 South-Main & South-Elm outage cases and neuron 21 classifies 3 out of 4 North-South & South-Elm outages.

2 The borders of the clusters are softened. For example neuron 14 now classifies the case of North-South & South-Elm at load level 110%. This corresponds to an intermediate load pattern where the load of line North-Lake gradually increases from minimal 115% at neuron 21, 163% at neuron 14 to maximal 214% at neuron 7. On the other hand, the load on line Main-Elm varies from maximal 71% at neuron 21 to 77% at neuron 14 down to 37% at neuron 7.

   Neuron 41 classifies some of the North-Lake outages thus linking clusters E and F. Neurons 11, 19 and 20 extend cluster E for the cases of 90% load.

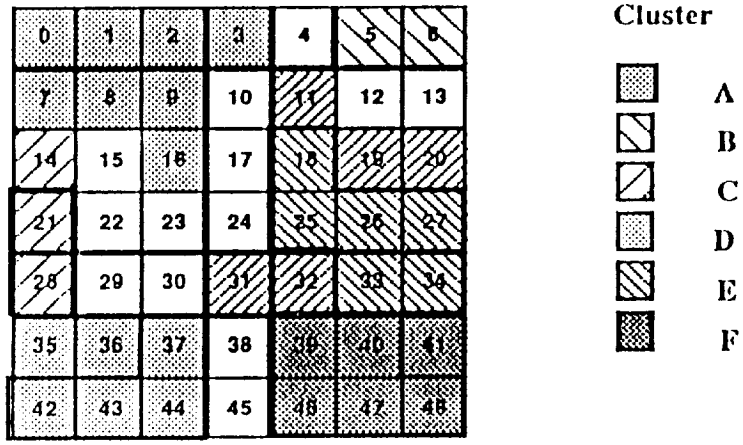
3 The cluster B classifies the special cases (isolated buses and total generation outage). It is still separated by "empty" neurons from its neighboring clusters. Clusters A, C and D are now linked. They as well as cluster B are still separated from clusters E and F since neurons situated on a diagonal like neuron 5 and neuron 11 are not neighbors of first order.

Table 3-5 shows the cluster map for the classification of vectors which have not been trained.

3.4 Future Work

Based on the work of Kohonen (1989) and Ritter (1988) the relation between number of clusters, number of neurons and size of the power system has to be investigated. Scaling issues for large size power systems have to be addressed. Like neural networks in general, this classification technique is supposed to be implemented ultimately in dedicated hardware in order to be used under real-time conditions.

Table 3-5: Cluster map for classification of 184 untrained vectors



3.5 Conclusions

This Section demonstrates the feasibility of classification of load patterns for power system static security assessment using the Kohonen self-organizing feature map. Our results are based on simulations of the Kohonen network performed on a work station and not on neural net hardware.

Therefore performance comparison with existing algorithms is not the issue of this paper. While the off-line training is rather time consuming the classification phase executes fast even on the simulator.

Instead emphasis is put on the interpretation of the obtained results. Simulation results show that indeed the self-organizing feature map projects safe, critical and unsafe operation points to different clusters. Knowing which cluster relates to which security state and mapping an operational point to one of these immediately identify whether the system is operating safely or not. Depending on the neural network size the mapping further splits up the safe, critical, and unsafe clusters into more specific subclusters corresponding to different load patterns and therefore to different regions in the security space.

The most important and original aspect of this neural net is its generalization property. The Kohonen map not only classifies successfully 184 loading patterns which have not been presented during the training phase but also is capable of extending clusters for vectors which do not fall in the same category as any trained one. This powerful and versatile feature is especially useful for power system operation where it is unrealistic to expect that all possible system states can be simulated off-line.

REFERENCES

Aggoune, M. E. et al., "Artificial Neural Networks for Power System Static Security Assessment", Proceedings of IEEE-ISCAS, Portland, 1989, pp. 490-494.

Chow, J.C. et al., "An Improved Hopfield Model for Power System Contingency Classification", Proceedings of IEEE-ISCAS, New Orleans, 1990, pp. 2925-2928.

Cottrell, M., Fort, J., "A Stochastic Model of Retinotopy: A Self-Organizing Process", Biol. Cybernetics no. 53, 1986, pp 405-411.

Dillon, T.S. et al., "Short Term Load Forecasting Using Adaptive Pattern Recognition and Self-Organizing Techniques", Proceedings of the 5th PSCC, Cambridge, 1975, pp 1-16.

Ejebe, G.C. and Wollenberg, B.F., "Automatic Contingency Selection", IEEE Trans. on Power, Apparatus and Systems, vol. PAS-98, no. 1, Jan/Feb 1979.

EPRI EL-2343, "Feasibility of Using Associative Memories for Static Security Assessment of Power System Overloads", investigator Y.-H. Pao, Project 1047-2, Final report, April 1982.

Fischl, R. et al., "The Application of Decision Theory to Contingency Selection", IEEE Trans. on Circuits and Systems, vol. CAS-29, no. 11, 1982, pp. 712-723.

Hochet, B., Peiris, V., Corbaz, G. and Declerq, M. "Implementation of a Neuron Dedicated to Kohonen Maps with Learning Capabilities", Proceedings of the IEEE CICC, Boston, 1990.

Hubel, D.H., and Wiesel, T.N, "Functional Architecture of Macaque Monkey Visual Cortex", Proceedings of the Royal Society, London, B, 1-59, 1977, pp 198.

Kohonen, T., "Self-Organized Formation of Topologically Correct Feature Maps", Biol. Cybernetics no. 43, pp 59-69, 1982.

Kohonen, T., "Self-Organization and Associative Memory", 3rd edition, Springer-Verlag, Berlin, 1989.

von der Malsburg, C., "Self-Organization of Orientation Sensitive Cells in the Striate Cortex", Kybernetik, 14, pp 85-100, 1973.

Oh, S. Y., "A Pattern Recognition and Associative Memory Approach to Power System Security Assessment", IEEE Trans. on Systems, Man and Cybernetics, vol. SMC-16, no. 1, Jan/Feb. 1986, pp. 62-72.

Pang, C. K., Prabhakara, F. S., El-Abiad, A. H. and Koivo, A. J., "Security Evaluation in Power Systems Using Pattern Recognition", IEEE Trans. on Power, Apparatus and Systems, vol. PAS-93 no. 2, May/June 74, pp. 969-976.

Ritter, H. and Schulten, K. "Convergence Properties of Kohonen's Topology Conserving Maps: Fluctuations, Stability and Dimension Selection", Biol. Cybernetics no. 60, 1988, pp 59-71.

Ritter H., and Schulten K., "Extending Kohonen's Self-Organizing Mapping Algorithm to Learn Ballistic Movements" in Eckmiller and von der Malsburg, Eds., "Neural Computers", Springer, Heidelberg, 1987, pp 393-406.

Sobajic, D. J. and Pao, Y.-H., "An Artificial Intelligence System for Power System Contingency Screening", Proceedings of the IEEE-PICA Conference, Montreal, May 87, pp. 107 ff.

Sobajic, D. J. and Pao, Y.-H., "Artificial Neural-Net Based Dynamic Security Assessment of Electric Power Systems", IEEE PES 1988 Winter Meeting, New York, Jan 31-Feb. 5, 1988, paper 88WM 211-5.

Stagg, G. and El-Abiad, A., "Computer Methods in Power System Analysis", McGraw-Hill, New York, 1968, pp 284.

Stott, B., "Review of Load-Flow Calculation Methods", Proceedings of the IEEE, vol. 62, no. 7, July 1974, pp. 916-929.

Tryba V., Metzen S. and Goser K., "Designing of Basic Integrated Circuits by Self-Organizing Feature Maps", Proceedings of Neuro-Nîmes, 1988.

Vittoz, E., Sorouchyari, E., Heim, P., Arreguit, X., and Krummenacher, F. "Analog VLSI Implementation of a Kohonen Map", Proceedings of the International Conference on Neural Networks, EPF-Lausanne, October 1989.

Wu, F. F. and Kumagai, S., "Steady-State Security Regions of Power Systems", IEEE Trans. on Circuit and Systems, vol. CAS-29, no. 11, no.v 1982, pp. 703 ff.

APPENDIX A

THEORY FOR A HIGH VOLTAGE SENSOR BASED ON ELECTRO-OPTIC POLYMER

The purpose of this section is to serve as a theoretical foundation for understanding the experimental work described in section 2. Section A.1 is a heuristic introduction to polarization and linear electro-optic effects. Section A.2 is an introduction to Jones calculus. In section A.3 the polarimetric configuration used in the sensor and test set up is discussed. In section A.4 the sensor configuration is analyzed by applying Jones calculus. In section A.5 the retardation phenomenon in the electro-optic polymer is analyzed based on the concept of an index ellipsoid. In section A.6 the above derived theory is used to describe the electro-optic response of a sample of poled electro-optic polymer, with numeric examples based on the measured data. It should be emphasized that the short introduction to the theory of polarization and electro-optic effect is only summary in its nature and is not intended to give the reader a complete understanding of these disciplines of optics.

A.1 Introduction to Polarization

Polarization of light, the interaction between polarimetric components and light, and the linear electro-optic effect are described. The description is mostly heuristic, but a short introduction to mathematical tools is also given.

A.1.1 The polarization of light

Classically light is described as electromagnetic waves. As such, light can be defined by two vectors **E** and **H**, called the electric and the magnetic field vectors. **E** and **H** are both perpendicular to the direction of propagation of the light wave.

By definition, the polarization of a light wave is specified by the electric field vector $\mathbf{E}(r,t)$ at a fixed point in space, r, at time t. The time variation of the electric field vector **E** of a monochromatic wave is sinusoidal. Assuming that the light is propagating in the z direction, the electric field vector will lie on the xy plane. Since the x component and the y component of the field vector can oscillate independently, one must consider the effect of these two oscillating orthogonal components.

Imagine that we let our principal system follow the light wave in time and space, and only watch the movement of the **E** field vector in the xy plane. What kind of curve would the end point of the electric field vector describe ?

If the x and y component are out of phase, the general solution to this problem is an ellipse. Now imagine gradually changing the phase between the x and the y component. Changing the phase would change the ellipticity of the ellipse. Special solutions for certain phase angles would be a circle and a straight line. Also the direction of rotation of the **E** field vector will change.

In describing polarized light, several categories will be discussed. The terminology will be summarized here:

Linearly polarized light The tip of the **E** vector moves along a straight line.

Elliptically polarized The tip of the **E** vector describes an ellipse.

Circularly polarized The tip of the **E** vector describes a circle.

Right handed polarization When viewed by an observer facing the approaching wave, the tip of the **E** field vector is seen to move counterclockwise.

Left handed polarization When viewed by an observer facing the approaching wave, the tip of the **E** field vector is seen to move clockwise.

Orthogonal polarization states Two perpendicular polarized states of linear polarized light are said to be orthogonal. The term orthogonal also covers the elliptical and circular states. Orthogonal circular states are counter rotating.

The superposition principle A given state of polarization can always be analyzed as the superposition of two orthogonal polarization states. The analyzed polarization state is then fully described by the two orthogonal states and the phase angle between them. The most common way the superposition principle is applied is when a given state of polarization is analyzed as the superposition of two linearly polarized polarization states, split up after the x and y directions in a coordinate system with the z axis being the direction of propagation of the light wave.

Polarization components

Polarization components are optical components (ie hardware) that influence the state of polarization of light. An easy way to understand polarization components is to apply the superposition principle and split up the incoming light wave in orthogonal linear states of polarization, parallel to characteristic directions of the components. The two orthogonal states will be influenced differently by the polarization component. After having passed the polarization component, the two orthogonal states can be imagined recombining into a new state of polarization that may be different from that of the incoming light wave.

A.1.2 The linear polarizer

A linear polarizer is shown in Figure A.1. It is characterized by transmitting linearly polarized light in a direction of polarization coinciding with a characteristic direction of the component, called the polarization axis. The component can be analyzed by splitting the incoming light into two linear polarized components, one parallel to the polarization axis, and the other perpendicular to it. Only the parallel state of polarization is transmitted, the other state is cancelled out. The linear polarizer converts any state of polarization into this linear state. However, the amplitude of the **E** field of the transmitted linear state is a function of the state of polarization of the incoming light wave.

A.1.3 The polarizing beamsplitter

This component can be understood as the combination of two linear polarizers, with their axes of polarization perpendicular to each other. Instead of transmitting a single linear state of polarization, it splits the incoming light wave up into two orthogonal states of polarization, separates them physically and transmits both of them. Because of the physical separation, the two polarization states do not recombine after having passed through the component. Consequently, their polarization states are preserved. This is shown in Figure A.2.

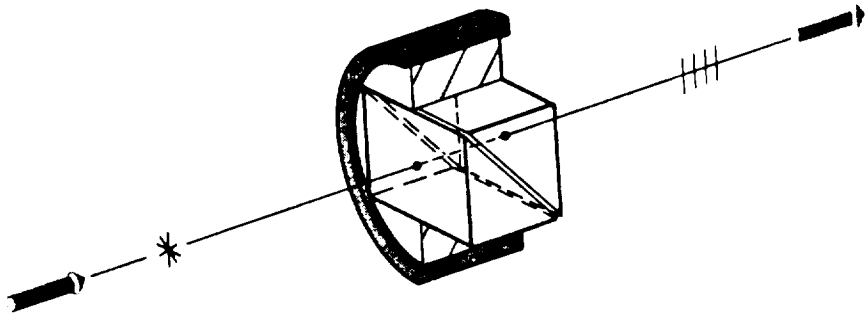


Figure A.1 Linear polarizer: Converts unpolarized light to linearly polarized light.

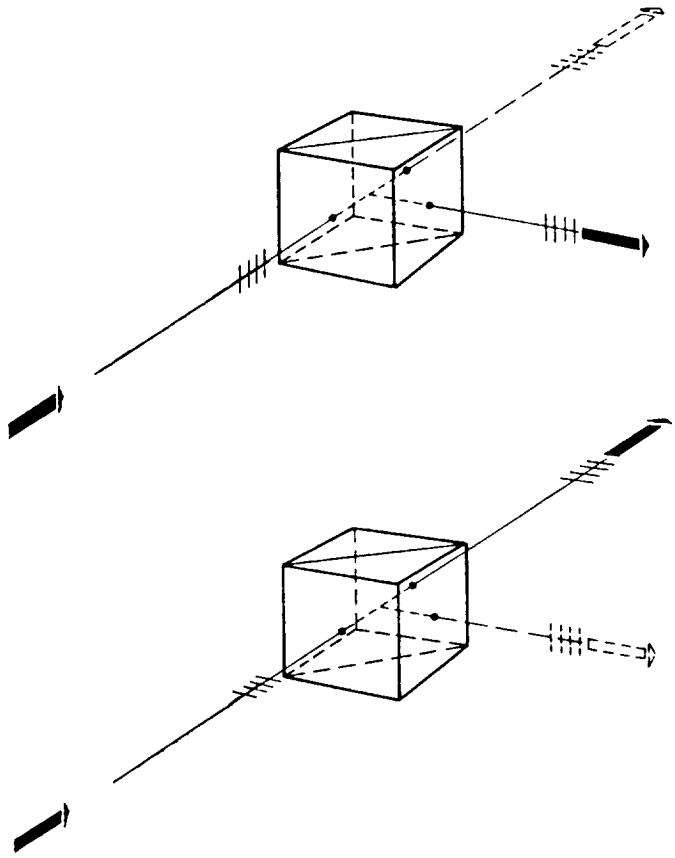


Figure A.2 Polarizing beamsplitter: Splits up an incoming beam into two beams with orthogonal linear polarizations.

A.1.4 Phase retarders

These components are often called wave plates. Examples are the 1/4 wave plate, the 1/2 wave plate and the Babinet-Soleil compensator, the last of which is an adjustable waveplate.

The function of a phase retarder is based on the phenomenon **linear birefringence.** A phase retarder has two characteristic axes, one of which is called the fast axis, the other of which is called the slow axis.

The component can be analyzed by imagining the incoming light wave split into two orthogonal linearly polarized states, one of which is parallel to the fast axis, while the other is parallel to the slow axis. The polarization state parallel to the fast axis will have a higher propagation speed than the one that is parallel to the slow axis. **This phenomenon is referred to as phase retardation or linear birefringence.**

Because of the difference in propagation time, a phase delay will be introduced between the two states after the light has passed the component. When the two polarization states recombine after having passed the phase retarder, the resulting polarization state is different from that of the incoming light wave. The polarization state of the transmitted light wave is completely defined by the polarization state of the incoming light wave and the phase delay introduced by the phase retarder. Phase delay can be positive as well as negative. Phase delay is additive, so the phase delay of a series of phase retarders is equal to the sum of the individual phase delays with their signs included. Phase retarders are often designed to introduce a phase delay of a certain fraction of a full wave. From this application comes the name wave plate. Normally the wave plates are characterized by the phase delay they introduce. For example, a wave plate which provides a phase delay of 1/4 wave is denoted a 1/4 wave plate. It should be noted that wave plates are usually wavelength specific devices, so a 1/4 wave plate would usually only provide an exact phase delay at the design wavelength. Figure A.3 shows a 1/4 wave plate.

A.1.5 The circular polarizer

The combination of a linear polarizer and a 1/4 wave plate is often used to generate a circularly polarized wave, therefore this combination is often referred to as a circular polarizer. Its mode of operation is explained below.

A beam of linearly polarized light is launched into a 1/4 wave plate with its direction of polarization at an angle of 45° to the fast and the slow axes.

The incoming linear light wave is split up into two linear polarized waves of equal amplitude. As the two waves propagate through the wave plate, a phase delay of a 1/4 wave occurs. When the two linear modes recombine after the having passed the 1/4 wave plate, the result is a circularly polarized wave because of the 1/4 wave phase delay.

A.1.6 The Babinet-Soleil Compensator

If the phase delay in the circular polarizer example above had been a different fraction of a wavelength, a different state of polarization would have been created. With a Babinet-Soleil compensator the phase delay is adjustable, and consequently any state of polarization can be created.

Conclusion: A 1/4 wave plate is able to convert linearly polarized light into circularly polarized light. A Babinet-Soleil compensator is able to convert linearly polarized light into any state of polarization. In fact it can convert any state of polarization into any other state of polarization.

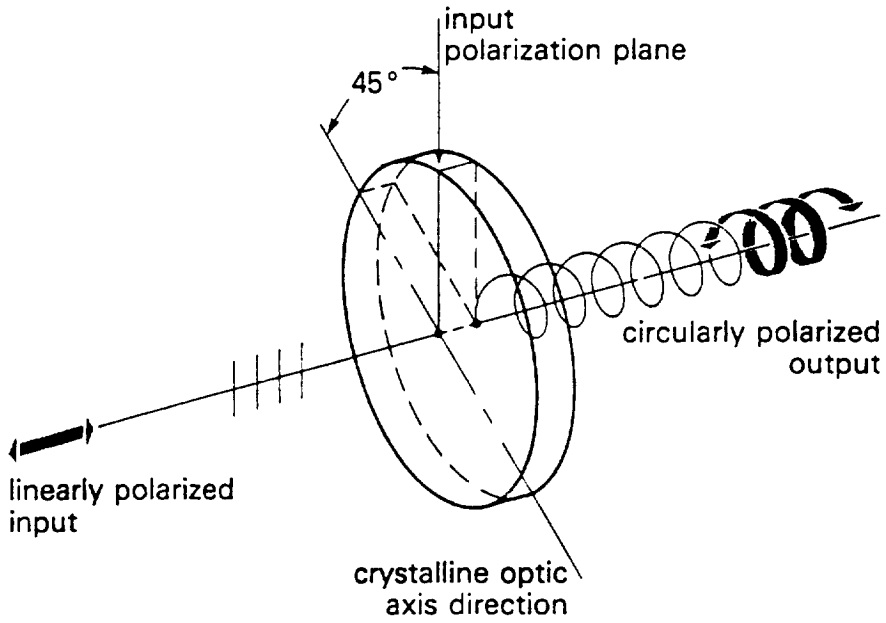


Figure A.3 1/4 wave plate: Can transform linearly polarized light into circularly polarized light, when the angle between the plane of linear polarization and the optic axis of the 1/4 wave plate is exactly 45 deg.

A.1.7 The electro-optic phase retarder

An electro-optic phase retarder induces a phase delay which is a linear function of the applied electric field. For an ideal device, the phase delay would be zero with no electric field applied, and directly proportional to the field when a field is applied. In practice, a certain amount of intrinsic birefringence will often be present. This means that the phase delay originating from the electric field will be superposed on a finite amount of constant phase delay. The phase delay thus is still a linear function of the electric field. For analysis, the component can be treated as an ideal electro-optic phase retarder followed by a regular phase retarder providing the constant phase delay. It may be necessary to compensate for undesired intrinsic birefringence by using not a 1/4 wave plate, but a wave plate that has a phase delay which, superposed on the intrinsic birefringence, will add up to a 1/4 wave phase delay.

The electro-optic phase retarder can be implemented in crystals or electro-optic polymer. In the case of electro-optic crystals the crystal axes have to be aligned correctly relative to the electric field, the propagation direction of the optical wave and the axes of polarizers and wave plates. In the case of the electro-optic polymer, it is the alignment of the poling axis which is important.

A.1.8 Detection of polarization

Photo detectors are not sensitive to the polarization of light, but to optical power. In order to detect polarization, it is necessary to let the light wave pass through a linear polarizer before it is incident on the detector. Used in this application, the linear polarizer is usually called an analyzer. Optical power is proportional to the square of the **E** field. Because the E field of the transmitted wave is a function of the state of polarization of the incoming light wave, the transmitted optical power also is a function of the state of polarization.

Conclusion: The analyzer is able to decode a change in polarization state into a change in optical power.

A.2 Introduction to Jones Calculus

The above description of polarization and polarization components has been more heuristic than strictly mathematical. It is important to get to a physical understanding before getting into the abstract mathematics. However, to perform a structured analysis of a polarimetric system, mathematical treatment is necessary. A number of mathematical tools have been developed to describe polarization. It is not the purpose of this report to describe them. However, the reader is urged to refer to the literature for more information.

The analysis of the amplitude modulator, which is the basis of the HV sensor, is performed with Jones calculus. A short introduction to complex-function notation and Jones calculus will be given.

Complex-function representation

In the complex-function notation representation, the electric field vector of a monochromatic plane wave propagating in the z direction is given by

$$\mathbf{E}(z,t) = \text{Re}\,[\mathbf{A}e^{j(\omega t - kz)}] \qquad (A.2.1)$$

where **A** is a complex vector which lies in the xy plane, ω is the angular frequency for this particular wavelength, k is the wave number and t is time.

The curve which the end point of the electric field vector describes in space is the time-evolution locus of the points whose coordinates (E_x, E_y) are

$$E_x = A_x \cos(\omega t - kz + \delta_x) \qquad (A.2.2)$$

$$E_y = A_y \cos(\omega t - kz + \delta_y) \qquad (A.2.3)$$

where δ_x and δ_y are phase constants and the complex vector **A** is defined as

$$\mathbf{A} = \hat{\mathbf{x}} A_x e^{j\delta_x} + \hat{\mathbf{y}} A_y e^{j\delta_y} \qquad (A.2.4)$$

where A_x and A_y are positive numbers and $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are unit vectors.

Jones calculus

In the calculus invented by R.C. Jones in 1940-41, the state of polarization of light is represented by a two-component complex vector, while each polarization component of a system is represented by a 2×2 complex matrix.

A matrix for a whole system is obtained by multiplying all the matrices, and the polarization state of the transmitted light is derived by multiplying the vector representing the input beam by the overall matrix. In the Jones calculus the optical system is treated as ideal, that is to say that the system is free from any loss.

The Jones vector formulation

According to the Jones vector concept the polarization state of a plane wave is expressed as a vector

$$\mathbf{V} = \begin{bmatrix} V_x \\ V_y \end{bmatrix} = \begin{bmatrix} A_x e^{j\delta_x} \\ A_y e^{j\delta_y} \end{bmatrix} \qquad (A.2.5)$$

The Jones vector is a complex vector. It is not a vector in real physical space, but in abstract mathematical space.

To obtain, as an example, the real x component of the electric field, the operation

$$E_x(t) = \text{ Re } [V_x e^{j\omega t}] = \text{ Re } [A_x e^{j(\omega t + \delta_x)}] \qquad (A.2.6)$$

must be performed.

The Jones vector contains complete information about the amplitude and the phases of the electric field vector components.

The Jones matrix formulation

The optical polarization components are represented in this formulation by complex 2×2 matrices. In general, a Jones matrix can be represented by

$$\mathbf{W} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad (A.2.7)$$

where a,b,c and d are complex numbers.

If $\mathbf{W}$ represents the Jones matrix of an optical system, and $\mathbf{V}$ represents the polarization state of the incoming wave, the polarization state of the transmitted wave can be represented by $\mathbf{V}'$, where

$$\mathbf{V}' = \mathbf{W} \cdot \mathbf{V} \qquad (A.2.8)$$

A.3 The Measurement Principle of the Sensor

The electro-optic high voltage sensor described in this report is often seen referred to in the literature as a Pockels cell. However the configuration is, from an optical point of view, identical to what is known as a transverse electro-optic amplitude modulator. The only difference is that, in the case of a sensor, the voltage represents the unknown parameter to be measured, while in the case of a modulator, the modulation voltage is known. An optical schematic representation of the high voltage sensor can be seen in Figure A.4.

A.3.1 The polarization of the system

The unpolarized light wave is converted into a vertically linearly polarized wave by the polarizer. The vertically polarized wave now enters the 1/4 wave plate, which separates the wave into two linearly polarized wave components of equal amplitude. One of the two wave components is parallel to the fast axis, which is denoted " f " in Figure A.4. The other wave component is parallel to the slow axis, which is denoted " s ". Because of the difference in propagation speed of the two wave components, a phase delay of 1/4 wave occurs. The phase delay makes the transmitted beam circularly polarized.

After having passed the 1/4 wave plate, the now circularly polarized light wave enters the electro-optic retarder. The electro-optic retarder is assumed to be ideal, that is to say it does not posses any intrinsic birefringence.

The electro-optic retarder is modulated with the electric field denoted E in Figure A.4. As the purpose of the sensor is to measure the line voltage on the 1 to 20 kV electric power distribution system, the electric field is assumed to be a 60-Hz ac signal. A phase retardation with a waveform identical to the waveform of the modulation field (i.e. the E field) is induced in the electro-optic retarder. If the waveform of the modulation field is a sinusoidal function of time, the waveform of the phase delay as a function of time will also be sinusoidal. As the phase delay changes from its negative peak value through zero to its positive peak value, it will modulate the ellipticity of the polarization of the transmitted light wave. When the induced phase delay is zero, the light remains circularly polarized. When the phase delay reaches its peak values, the polarization state of the light will be elliptical. The ellipticity of the two elliptical states will be the same from a geometric point of view. However, in one state the ellipse will be standing up, in the other it will be lying down. The transmitted wave will alternate around a circularly polarized state between two elliptically polarized states, one of them being represented by a tall and slim ellipse the other by a short and fat ellipse.

After having passed the electro-optic retarder, the alternating wave enters the analyzer. The analyzer is placed so that its polarization axes are parallel to the two axes of deformation of the ellipse of the elliptically polarized light. A polarizing beam splitter has been chosen in this analysis solely to demonstrate that there are two angles at which the analysis can be performed.

The alternation of the elliptically polarized light modulates the optical power of the two, orthogonally transmitted, linearly polarized light waves. **In this way the waveform of the modulating electric field is superposed as a modulation on top of the transmitted optical power.** However, one of the two signals transmitted by the polarizing beamsplitter is in phase with the modulation field, the other is in anti-phase.

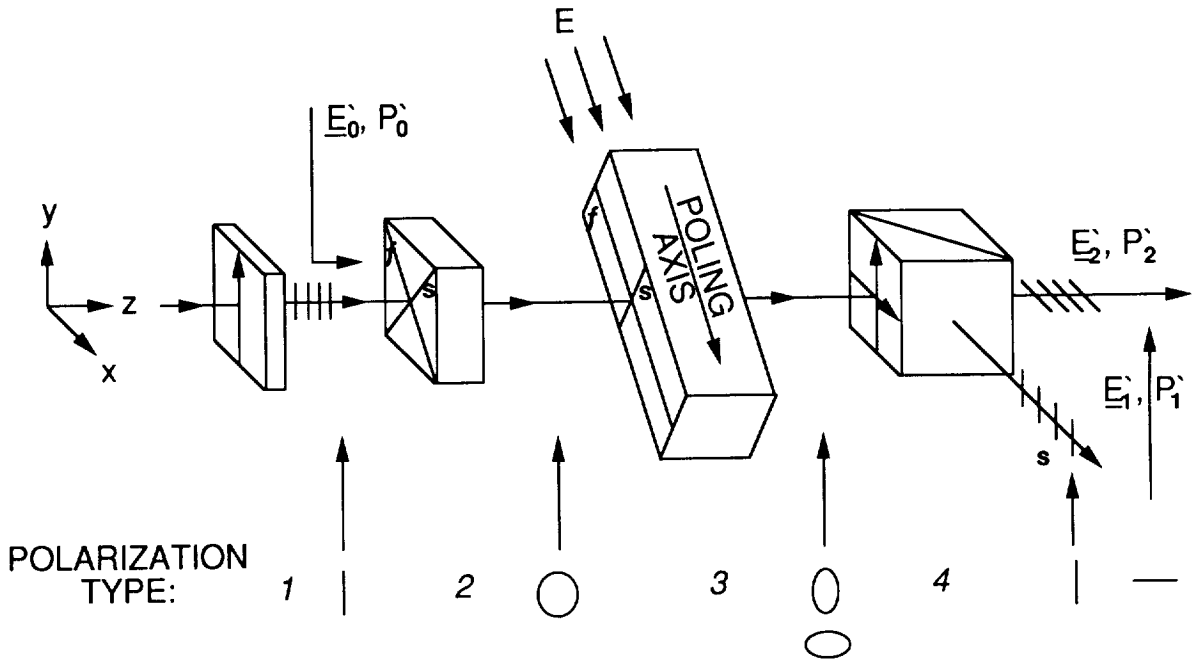


Figure A.4 Principle of electro-optic voltage sensor.
Furthest left are shown the principal axes of the chosen coordinate system. A wave of unpolarized light represented by an arrow is launched into the system from the left towards the right:

(1) represents a linear polarizer, here called the polarizer.
(2) represents a 1/4 waveplate.
(3) represents a transversely modulated electro-optic retarder, here called the electro-optic retarder.
(4) represents a polarizing beamsplitter used as an analyzer, here called the analyzer.

A.3.2 Deriving the signal from the transmitted optical power

By applying Jones calculus analysis to the system shown in Figure A.4, it will be shown in Section A.4 that the transmitted optical power is given by

$$P_1' = \frac{1}{2}P_0'(1 - \sin[\Gamma(t)]) \qquad (A.3.1)$$

$$P_2' = \frac{1}{2}P_0'(1 + \sin[\Gamma(t)]) \qquad (A.3.2)$$

where $\Gamma(t) = kE(t)$, P_1', P_2' are the transmitted optical power in channels 1 and 2 and P_0' is the transmitted optical power after the light wave has passed the polarizer (see Figure A.4). If the incident light beam is completely unpolarized, P_0' will be exactly half of the power of the incident light wave. $\Gamma(t)$ is the electro-optically induced phase retardation as function of time. $E(t)$ is the electric modulation field as function of time, and k is a constant.

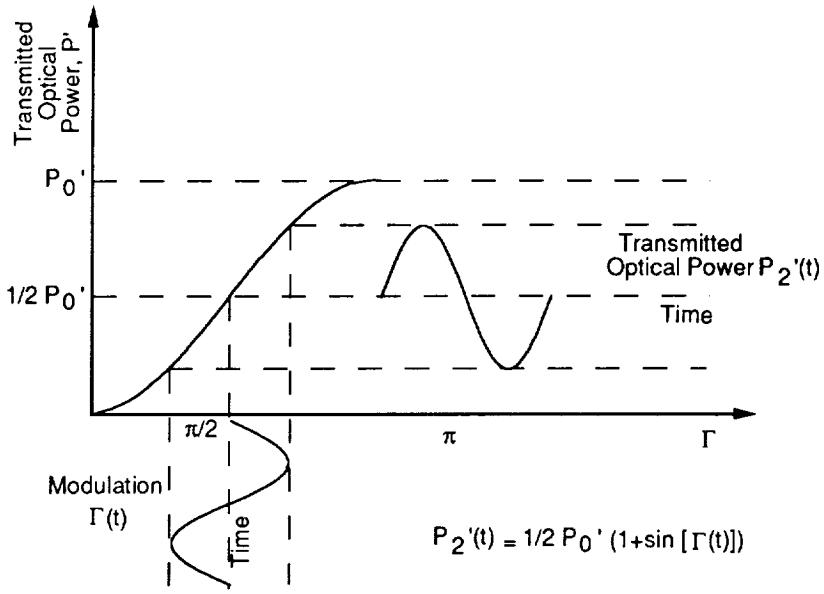


Figure A.5 Transmitted optical power as function of retardation. Electro-optic modulation at $\Gamma = \pi/2$ is exemplified, $\Gamma(t)$ being the modulation as a time function and $P_2'(t)$ the transmitted optical power as a function of time.

Consider the signal

$$P_2' = \frac{1}{2}P_0'(1 + \sin[\Gamma(t)]) \tag{A.3.3}$$

The signal can be separated into a dc part and an ac part:

$$P_{ac}' = \frac{1}{2}P_0' \sin[\Gamma(t)] \tag{A.3.4}$$

$$P_{dc}' = \frac{1}{2}P_0' \tag{A.3.5}$$

The ratio of the ac part to the dc part determines the voltage, and is independent of the optical power.

This ratio is referred to as the modulation index $m(t)$:

$$m(t) = \frac{P_{ac}'}{P_{dc}'} = \sin[\Gamma(t)] = \sin[kE(t)] \tag{A.3.6}$$

For sensor applications a very small modulation index is typically used, so the approximation $\sin x = x$ is valid.

In this case, it can be approximated that

$$m(t) = \frac{P_{ac}'}{P_{dc}'} = \Gamma(t) = kE(t) \tag{A.3.7}$$

If only one channel is available $E(t)$ can be determined from equation (A.3.7). If both channels are available, $\Gamma(t)$ can be derived from the difference/sum relation

$$\frac{P_2' - P_1'}{P_2' + P_1'} = \Gamma(t) = kE(t) \qquad (A.3.8)$$

With two channels, both the dc and the ac components can be detected. For a real device, the optical power P_0 injected into the system from the optical source is more appropriate than the entity P_0'. In an ideal system with no loss, and assuming a totally unpolarized wave entering the polarizer, the relation between P_0' and P_0 is determined by a 50% polarization loss, yielding

$$P_0' = \frac{1}{2}P_0 \qquad (A.3.9)$$

In a real system, loss will be present. Assume that the losses in the two channels are given by the attenuation coefficients a and b. The attenuation coefficients represent loss in the transmitter and detector fibers, coupling loss, reflection loss and also the 50% polarization loss that occurs when the unpolarized wave passes through the polarizer. As a consequence, it is given that $a, b < 0.5$.

For a real system, the detected optical power in the two channels is given by

$$P_1 = \frac{1}{2}aP_0\big(1 - \sin[\Gamma(t)]\big) \qquad (A.3.10)$$

$$P_2 = \frac{1}{2}bP_0\big(1 + \sin[\Gamma(t)]\big) \qquad (A.3.11)$$

For each signal an ac component and a dc component can be defined. For the signal P_2, the components are:

$$P_{ac} = \frac{1}{2}bP_0 \sin[\Gamma(t)] \qquad (A.3.12)$$

$$P_{dc} = \frac{1}{2}bP_0 \qquad (A.3.13)$$

It can be seen that the relation

$$m(t) = \frac{P_{ac}}{P_{dc}} = \sin[\Gamma(t)] = \sin[kE(t)] \qquad (A.3.14)$$

is valid. The difference/sum relation is only valid if a and b are identical, which is generally not the case. However, it is possible electronically to compensate for this by providing different amplification for the two channels in such a way that their resulting dc components are equal.

A.4 Jones Calculus Analysis of the Sensor

Consider the optical system seen in Figure A.4. The system will now be analyzed utilizing Jones calculus. Each optical element can be represented by a matrix, and the light wave can be represented by a vector. The matrices and vectors are defined in the basic coordinate system with principal axes x and y. However, it is also convenient in the calculations to operate with a rotated coordinate system with principal axes s and f. The principal axes s and f, seen in Figure A.4 are the *slow* and *fast* axes of the 1/4 wave plate and the electro-optic retarder.

A.4.1 The polarizer and the incoming wave

The polarizer can be represented by the matrix

$$\mathbf{W}_p = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{A.4.1}$$

We cannot use this matrix directly in the calculations, because the incoming light on the polarizer is non-polarized, and consequently cannot be represented by a Jones vector. However, we know that the wave transmitted through the polarizer will be linearly polarized with its axis of polarization coinciding with the axis of polarization of the linear polarizer, and consequently can be represented by the vector

$$\mathbf{V}_0' = \begin{bmatrix} 0 \\ E_0' \end{bmatrix} \tag{A.4.2}$$

where E_0' represents the vertical electric field component. V_0' is related to the optical power P_0' transmitted through the polarizer through the equation

$$P_0' = \frac{c\varepsilon_0}{2}\mathbf{V}_0'^{*}\mathbf{V}_0' = \frac{c\varepsilon_0}{2}(E_{0'})^2 \tag{A.4.3}$$

where the asterisk represents complex conjugation.

A.4.2 The 1/4 wave plate and the electro-optic retarder

Now let us consider the rotated coordinate system *sf*. Letting (rot) indicate the rotated coordinate system the general expression for a retarder with its fast and slow axis coinciding with the f and s axis, is given by the matrix

$$\mathbf{W}_{r(rot)abs} = e^{-j\phi}\begin{bmatrix} e^{-j\frac{\Gamma}{2}} & 0 \\ 0 & e^{j\frac{\Gamma}{2}} \end{bmatrix} \tag{A.4.4}$$

where Γ is the phase retardation given by

$$\Gamma = (n_s - n_f)\frac{\omega l}{c} \tag{A.4.5}$$

where ω is the angular frequency of the monochromatic wave, c is the speed of light in a vacuum and n_f and n_s are the indices of refraction along the fast and slow axes of the device. The mean absolute phase change ϕ is given by

$$\phi = \frac{1}{2}(n_s + n_f)\frac{\omega l}{c} \tag{A.4.6}$$

As the absolute phase change is without importance when considering polarimetric measurements, this term will be omitted, and the matrix will only be represented by

$$\mathbf{W}_{r(rot)} = \begin{bmatrix} e^{-j\frac{\Gamma}{2}} & 0 \\ 0 & e^{j\frac{\Gamma}{2}} \end{bmatrix} \tag{A.4.7}$$

The 1/4 wave plate provides for a phase retardation of 1/4 wave, equal to $\pi/2$ radians. Consequently, the 1/4 wave plate can be represented in the rotated system by the matrix

$$\mathbf{W}_{qw(rot)} = \begin{bmatrix} e^{-j\frac{\pi}{4}} & 0 \\ 0 & e^{j\frac{\pi}{4}} \end{bmatrix} \tag{A.4.8}$$

The electro-optic retarder provides a phase retardation which is a function of the voltage, and consequently a function of time. Thus, the electro-optic retarder is represented in the rotated system by the matrix

$$\mathbf{W}_{er(rot)} = \begin{bmatrix} e^{-j\frac{\Gamma(t)}{2}} & 0 \\ 0 & e^{j\frac{\Gamma(t)}{2}} \end{bmatrix} \tag{A.4.9}$$

The combination of the two elements is, in the rotated system, represented by the matrix product of $\mathbf{W}_{qw(rot)}$ and $\mathbf{W}_{er(rot)}$. The resulting matrix is given by

$$\mathbf{W}_{(qw+er)(rot)} = \mathbf{W}_{er(rot)}\mathbf{W}_{qw(rot)} = \begin{bmatrix} e^{-j[\frac{\Gamma(t)}{2}+\frac{\pi}{4}]} & 0 \\ 0 & e^{j[\frac{\Gamma(t)}{2}+\frac{\pi}{4}]} \end{bmatrix} \tag{A.4.10}$$

A.4.3 Transformation between coordinate systems

In order to transform the matrix $\mathbf{W}_{(rot)}$ from the rotated coordinate system to the basic x,y system, where it is denoted $\mathbf{W}$, a transformation is necessary.

The transformation is given by

$$\mathbf{W} = \mathbf{R}(-\psi) \cdot \mathbf{W}(rot) \cdot \mathbf{R}(\psi) \tag{A.4.11}$$

where

$$\mathbf{R}(\psi) = \begin{bmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{bmatrix} \tag{A.4.12}$$

$$\mathbf{R}(-\psi) = \begin{bmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{bmatrix} \tag{A.4.13}$$

and ψ is the rotational angle between the two coordinate systems.

In the above case ψ is equal to $\pi/4$, and the two transformation matrices are given by

$$\mathbf{R}(\frac{\pi}{4}) = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \tag{A.4.14}$$

and

$$\mathbf{R}(-\frac{\pi}{4}) = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \tag{A.4.15}$$

The combination of the 1/4 wave plate and the electro-optic retarder is then given by the matrix $\mathbf{W}$, where

$$\begin{aligned} \mathbf{W}_{(qw+er)} &= \mathbf{R}(-\frac{\pi}{4})\mathbf{W}_{(qw+er)(rot)}\mathbf{R}(\frac{\pi}{4}) \\ &= \begin{bmatrix} \cos\left[\frac{\Gamma(t)}{2} + \frac{\pi}{4}\right] & -j\sin\left[\frac{\Gamma(t)}{2} + \frac{\pi}{4}\right] \\ -j\sin\left[\frac{\Gamma(t)}{2} + \frac{\pi}{4}\right] & \cos\left[\frac{\Gamma(t)}{2} + \frac{\pi}{4}\right] \end{bmatrix} \end{aligned} \tag{A.4.16}$$

A.4.4 The analyzer

In Figure A.4 a polarizing beamsplitter is shown as the analyzer, though a conventional linear polarizer is actually used in the practical implementation of the high voltage sensor. The reason for choosing the polarizing beamsplitter in this example is that it demonstrates the two orthogonal orientations possible for the axis of the analyzer.

The polarizing beamsplitter separates the light wave into two physically separated wave components. Consequently two Jones matrices are necessary to describe this optical element. One matrix is identical to that of a linear polarizer with the axis of polarization in the vertical plane, the other to a linear polarizer with the axis of polarization in the horizontal plane.

The matrix related to the vertical polarization is given by

$$\mathbf{W}_{1a} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{A.4.17}$$

The matrix related to the horizontal polarization is given by

$$\mathbf{W}_{2a} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \tag{A.4.18}$$

The total Jones matrix for each of the two channels can be given by multiplication with the Jones matrix for the retardation components. The matrices are given by

$$\mathbf{W}_1 = \mathbf{W}_{1a} \cdot \mathbf{W}_{qw+er} = \begin{bmatrix} 0 & 0 \\ -j\sin\left[\frac{\Gamma(t)}{2} + \frac{\pi}{4}\right] & \cos\left[\frac{\Gamma(t)}{2} + \frac{\pi}{4}\right] \end{bmatrix} \tag{A.4.19}$$

$$\mathbf{W}_2 = \mathbf{W}_{2a} \cdot \mathbf{W}_{qw+er} = \begin{bmatrix} \cos\left[\frac{\Gamma(t)}{2} + \frac{\pi}{4}\right] & -j\sin\left[\frac{\Gamma(t)}{2} + \frac{\pi}{4}\right] \\ 0 & 0 \end{bmatrix} \tag{A.4.20}$$

A.4.5 Output signals

The Jones vectors for the two transmitted beams can now be determined by multiplication with the Jones vector for the incoming wave after the polarizer. The two resulting vectors are given by

$$\mathbf{V}_1' = \mathbf{W}_1 \cdot \mathbf{V}_0' = \begin{bmatrix} 0 \\ E_0' \cos\left[\frac{\Gamma(t)}{2} + \frac{\pi}{4}\right] \end{bmatrix} \tag{A.4.21}$$

$$\mathbf{V}_2' = \mathbf{W}_2 \cdot \mathbf{V}_0' = \begin{bmatrix} -jE_0' \sin\left[\frac{\Gamma(t)}{2} + \frac{\pi}{4}\right] \\ 0 \end{bmatrix} \tag{A.4.22}$$

The transmitted power in the two channels is given by

$$P_1' = \frac{c\varepsilon_0}{2}\mathbf{V}_1'^{*} \cdot \mathbf{V}_1' = \frac{c\varepsilon_0}{2}(E_0')^2 \cos^2\left[\frac{\Gamma(t)}{2} + \frac{\pi}{4}\right] \tag{A.4.23}$$

$$P_2' = \frac{c\varepsilon_0}{2}\mathbf{V}_2'^{*} \cdot \mathbf{V}_2' = \frac{c\varepsilon_0}{2}(E_0')^2 \sin^2\left[\frac{\Gamma(t)}{2} + \frac{\pi}{4}\right] \tag{A.4.24}$$

This implies that there is a quadratic relationship between the optical power P_1', P_2' and the electric field E_0' of the wave. Using the relation between the incoming optical power P_0' after the polarizer and the electric field vector component E_0', the transmitted optical power P_1', P_2' in the two channels can be represented by

$$P_1' = P_0' \cos^2\left[\frac{\Gamma(t)}{2} + \frac{\pi}{4}\right] = \frac{1}{2}P_0'(1 - \sin[\Gamma(t)]) \tag{A.4.25}$$

$$P_2' = P_0' \sin^2\left[\frac{\Gamma(t)}{2} + \frac{\pi}{4}\right] = \frac{1}{2}P_0'(1 + \sin[\Gamma(t)]) \tag{A.4.26}$$

These two expressions are the well known equations for the electro-optic amplitude modulator, which were introduced in the previous section A.3.

A.5 Phase Retardation in Electro-optic Polymer

In subsection A.3.2, the electrically induced phase retardation was introduced as a time function $\Gamma(t)$, and it was claimed that the relationship between $\Gamma(t)$ and the electric modulation field $E(t)$ is given by

$$\Gamma(t) = kE(t) \tag{A.5.1}$$

The fundamental purpose of this subsection is to determine the nature of the constant k. k is a function of the geometrical alignment of the poling axis of the polymer relative to the direction of the electric modulation field and the direction of propagation of the optical wave. However, k is also related to the molecular properties of the electro-optic dye, to the success of the curing process of the polymer and to the field strength of the electric poling field.

In order to determine the relationship between $\Gamma(t)$ and $E(t)$ it is necessary first to go through quite elaborate theoretical considerations.

A.5.1 Linear electro-optic effect

The electrically induced phase retardation $\Gamma(t)$ described above is caused by the linear electro-optic effect, which will be explained below.

The propagation of polarized light waves in anisotropic media can be described by the index ellipsoid surface. The concept of the index ellipsoid is known from optics theory. In general the electro-optic effect is a change in both the dimensions and the orientation of the index ellipsoid.

The electro-optic effect in organic materials is thought to be electronic in its origin and results from the redistribution of charge in the molecule due to the application of an external electric field. Electro-optic effects can be grouped into two categories. These are the linear electro-optic effect (the electro-optic effect is a linear function of the applied external electric field), also called Pockels' effect after F. Pockels, who first studied it in 1893; and the quadratic electro-optic effect (the electro-optic effect is a quadratic function of the applied external electric field), also called Kerr effect after J. Kerr, who first studied the effect in 1875.

In most practical applications of the electro-optic effect, the applied electric field is small compared to the electric field inside the atom, which is typically in the order of 10^8V/cm. As a result, the quadratic effect is in most cases negligible compared to the linear effect. However for some materials with centro-symmetric point groups, the linear electro-optic effect vanishes because of symmetry considerations and the quadratic effect becomes the dominant phenomenon.

Poled electro-optic polymer is non-centrosymmetric, and therefore the linear electro-optic effect is finite or nonvanishing, and much larger than the quadratic effect for the relatively weak electric modulation fields which are applied.

The electro-optic effect is defined through a change in the optical impermeability tensor by an applied electric field. For the linear effect, this change is given by

$$B_{ij}(E) - B_{ij}(0) = \left[\frac{1}{\epsilon(E)}\right]_{ij} - \left[\frac{1}{\epsilon(0)}\right]_{ij} \equiv r_{ij,k}E_k \qquad (A.5.2)$$

where B_{ij} is the impermeability tensor, ϵ is the dielectric tensor, E_k is the applied electric modulation field, and $r_{ij,k}$ is the linear electro-optic coefficient. $r_{ij,k}$ is symmetric with respect to the two first indices and therefore can be written in contracted notation, $r_{u,k}$.

The linear electro-optic coefficients in the contracted notation $r_{u,k}$ are the elements in the 3×6 electro-optic tensor, which provides the information about the change in the index ellipsoid when an external electric field is applied. The tensor in its general form can be represented by

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ \vdots & \vdots & \vdots \\ r_{61} & r_{62} & r_{63} \end{bmatrix} \qquad (A.5.3)$$

However, in most cases many of the elements of the tensor are zero, and some of the elements may be identical due to symmetry considerations.

A.5.2 Macroscopic and molecular polarization

The next step is to determine the electro-optic tensor for the electro-optic polymer. In order to identify the elements of the tensor, it is necessary first to study the relationship between the linear electro-optic coefficients and the macroscopic polarization of electro-optic polymer.

From a macroscopic point of view the origin of nonlinear optical effects, such as the linear electro-optic effect, can be found in the expansion of the material polarization density in powers of the electric field, which in the electric-dipole approximation is given by

$$P_j = P_j^{(0)}(t) + \chi_{ij}^{(1)}(t)E_j(t) + \chi_{ijk}^{(2)}(t)E_j(t)E_k(t) + \chi_{ijkl}^{(3)}(t)E_j(t)E_k(t)E_l(t) + \ldots \quad (A.5.4)$$

The material polarization density expresses the separation of charge in the material. It quantifies the electric dipole-moment in the material (charge $\times$ distance).

The first two terms correspond to the spontanous polarization and to linear optical effects, respectively. The terms in higher orders of the electric field yield the various phenomena of nonlinear optics. The lowest order nonlinear term $\chi_{ijk}^{(2)}$, also called the second order susceptibility, leads to three-wave mixing and the linear electro-optic effect. The linear electro-optic effect is categorized as a nonlinear optical effect, because it is related to $\chi_{ijk}^{(2)}$.

The macroscopic second-order susceptibilities of poled electro-optic polymer are directly traceable to the non-linear properties of the electro-optic dye molecules.

In analogy to the macroscopic polarization of the electro-optic polymer, the molecular polarization of the electro-optic dye is given by

$$P_I(t) = \mu_I^0 + \alpha_{IJ}(t)E_J(t) + \beta_{IJK}(t)E_K(t) + \gamma_{IJKL}(t)E_J(t)E_K(t)E_L(t) + \ldots \quad (A.5.5)$$

where μ_I^0 is the molecular ground-state dipole moment, $\alpha_{IJ}(t)$ the linear polarizability, and $\beta_{IJK}(t)$ and $\gamma_{IJKL}(t)$ the two lowest-order nonlinear-optical polarizabilities (or hyperpolarizabilities). The scalar product of the vector μ_I^0 and the tensor $\beta_{IJK}(t)$ is an entity often referred to. In order to be in accordance with the simplified terminology often used, the indices will be suppressed and the scalar product will be referred to as the $\mu \cdot \beta$ product.

A.5.3 Point group symmetry

The poled electro-optic polymer is formed by incorporating the dipolar non-linear optical dye molecules into the resins of the epoxy. While the epoxy is in a state which allows the dye molecules high mobility, the dipolar molecules are oriented by the applied external electric poling field. The epoxy is then taken to a state of mechanical rigidity where the dye mobility is frozen out. The curing process results in an increase in molecular weight of the polymer. As the degree of curing increases, the glass transition temperature, T_g, of the epoxy increases. Ultimately, the epoxy is completely cured and the limiting T_g is obtained. The epoxy is cooled down to room temperature with the field still applied to maintain the induced alignment. Once equilibrated at room temperature the poling field is turned off and a net alignment in the direction of the field remains.

As a result of the poling process, the electro-optic polymer possesses only one unique symmetry axis, the poling axis defined by the electric poling field, in the direction of which the dipolar nonlinear-optical dye molecules are oriented.

This symmetry characteristic makes poled electro-optic polymer belong to the non-centrosymmetric point group ∞mm.

For poled electro-optic materials there exists the relationship

$$r_{ij,k} = -\frac{2}{n^2}\chi^{(2)}_{ijk} \tag{A.5.6}$$

between the linear electro-optic coefficient and the second order susceptibility, where n is the index of refraction.

For point group ∞mm the nonzero tensor components of the second-order macroscopic susceptibilities are

$$\chi^{(2)}_{333}, \chi^{(2)}_{113}, \chi^{(2)}_{223}, \chi^{(2)}_{131}, \chi^{(2)}_{232}, \chi^{(2)}_{311}, \chi^{(2)}_{322} \tag{A.5.7}$$

where

$$\chi^{(2)}_{333}, \chi^{(2)}_{113} = \chi^{(2)}_{223} \quad \text{and} \quad \chi^{(2)}_{131} = \chi^{(2)}_{232} \quad \text{and} \quad \chi^{(2)}_{311} = \chi^{(2)}_{322} \tag{A.5.8}$$

As a result of (A.5.6) the nonzero linear electro-optic coefficients are the coefficients with the same subscript, namely

$$r_{333}, r_{113}, r_{223}, r_{131}, r_{232}, r_{311}, r_{322} \tag{A.5.9}$$

Because of permutation symmetries in the indices i and j, the electro-optic coefficients are normally written in a contracted notation, where the indices ij are abbreviated to the index u after the rules defined below.

$$\begin{aligned} r_{1k} &= r_{11k}, \\ r_{2k} &= r_{22k}, \\ r_{3k} &= r_{33k}, \\ r_{4k} &= r_{32k} = r_{32k}, \\ r_{5k} &= r_{13k} = r_{31k}, \\ r_{6k} &= r_{12k} = r_{21k}, \\ k &= 1,2,3 \end{aligned} \tag{A.5.10}$$

The nonzero coefficients in the contracted notation are

$$r_{33}, r_{13}, r_{23}, r_{42}, r_{51} \tag{A.5.11}$$

where

$$r_{13} = r_{23} = r_{42} = r_{51} \tag{A.5.12}$$

and

$$r_{13} = \frac{1}{3}r_{33} \tag{A.5.13}$$

A.5.4 The electro-optic tensor

The electro-optic coefficients r_{uk} are elements in the electro-optic tensor. In general the electro-optic tensor for point group ∞mm is given by

$$\begin{bmatrix} 0 & 0 & r_{13} \\ 0 & 0 & r_{23} \\ 0 & 0 & r_{33} \\ 0 & r_{42} & 0 \\ r_{51} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{A.5.14}$$

By utilizing the relations between the tensor coefficients (A.5.12), the tensor can be simplified to

$$\begin{bmatrix} 0 & 0 & r_{13} \\ 0 & 0 & r_{13} \\ 0 & 0 & r_{33} \\ 0 & r_{13} & 0 \\ r_{13} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{A.5.15}$$

A.5.5 Electro-optic coefficients and molecular properties

Assuming that the molecular non-linear properties are not severely perturbed by neighboring molecules, the macroscopic susceptibility can be calculated as the statistical average of the molecular hyperpolarizabilities, and is represented by

$$\chi_{ijk} = N\langle\beta^*_{IJK}\rangle_{ijk} \cdot F \tag{A.5.16}$$

where N is the number density of the nonlinear-optical dye molecules, β^*_{IJK} the local field-corrected molecular second-order polarizability, the angle brackets denote the statistical average and F is a product of local field factors.

For point group ∞mm case, it can be shown that

$$r_{33} \sim N\mu\beta\frac{E_p}{5k_BT} \tag{A.5.17}$$

where the subscripts have been suppressed for simplicity, and

$$r_{13} = \frac{1}{3}r_{33} \tag{A.5.18}$$

where $\mu\cdot\beta$ is the simplified notation for the scalar product of the molecular dipole-moment vector μ^0_I and the second order polarizability tensor $\beta_{IJK(t)}$ for the dye. E_p is the field strength of the electric poling field, T is the absolute temperature at which the state transition of the polymer takes place, N is the number density of the dye molecules and k_B is Boltzmann's constant.

It should be noted that the number density N of the dye molecules is proportional to the concentration c of the dye in the compound of dye and polymer. As the dye concentration is simpler to operate with in experimental work, (A.5.17) will sometimes be written in the form

$$r_{33} \sim c\mu\beta\frac{E_p}{5k_BT} \qquad (A.5.19)$$

A.5.6 The index ellipsoid

For solving the index ellipsoid problem, we assume that the poled electro-optic polymer is isotropic when there is no external electric field present. That is to say, it has no intrinsic birefringence (same index of refraction in all directions). In this case, the index ellipsoid will reduce to a sphere represented by the following equation when there is no external electric field present

$$\frac{x^2}{n_0^2}+\frac{y^2}{n_0^2}+\frac{z^2}{n_0^2}=1 \qquad (A.5.21)$$

When an electric field is applied, the index ellipsoid will change both the dimensions and the orientation. In the presence of an electric field $\mathbf{E}=(E_x,E_y,E_z)$ the index ellipsoid will in general form be represented by the equation

$$(\frac{1}{n_0^2}+r_{1k}E_k)x^2+(\frac{1}{n_0^2}+r_{2k}E_k)y^2+(\frac{1}{n_0^2}+r_{3k}E_k)z^2+2yzr_{4k}E_k+2xzr_{5k}E_k+2xyr_{6k}E_k=1 \qquad (A.5.22)$$

where $E_k(k=1,2,3)$ is a component of the electric field, and 1,2,3 correspond to the principal axes x, y, z.

When substituting r_{uk} by the electro-optic tensor coefficients for point group ∞mm and E_1, E_2, E_3 with E_x, E_y, E_x, one gets the specific equation for the index ellipsoid in the poled electro-optic polymer when an electric field is present. The index ellipsoid is represented by

$$(\frac{1}{n_0^2}+r_{13}E_z)x^2+(\frac{1}{n_0^2}+r_{13}E_z)y^2+(\frac{1}{n_0^2}+r_{33}E_z)z^2+2r_{13}\Big[yzE_y+zxE_x\Big]=1 \qquad (A.5.23)$$

where the poling axis is here represented by the z axis.

In order to calculate the induced phase retardation, it is necessary to bring the ellipsoid to the form

$$\frac{x'^2}{n_{x'}^2}+\frac{y'^2}{n_{y'}^2}+\frac{z'^2}{n_{z'}^2}=1 \qquad (A.5.24)$$

where $n_{x'}$, $n_{y'}$, $n_{z'}$ are the new electric field dependent indices of refraction, and x', y', and z' are the new principal axes. The principal axes do not necessarily change. One or more of the axes x', y', and z' could be the same as x, y, and z.

The index ellipsoid is used as a tool to determine the indices of refraction for two orthogonal linear polarization states. Only if the difference of the two indices is a function of the applied electric modulation field will electro-optically induced phase retardation exist in the material.

In order to perform the analysis it is necessary to specify the direction of the electric modulation field, the direction of propagation of the light wave and calculate the electro-optically induced phase retardation for this particular case. In order to perform a thorough analysis of the material, several cases must be studied.

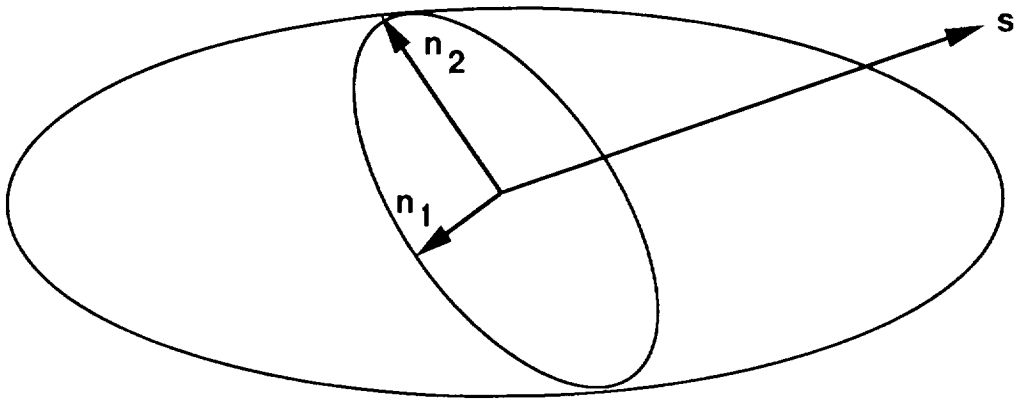


Figure A.6 Intersection between index ellipsoid and plane normal to the direction of propagation of the optical wave.

Imagine the index ellipsoid cut in half along a plane which is perpendicular to the direction of propagation of the light wave, and which goes through the center of the ellipsoid. The intersection of the ellipsoid and the plane is an ellipse, as shown in Figure A.6. The two indices of refraction are given by the two half axes of this ellipse. If linear electro-optic effect is present, the length of one or both of the axes is a linear function of the modulation field. In order to measure the induced phase retardation in the arrangement described in section A.3.1, the analyzer should be positioned with its polarization axis at an angle of 45 degrees relative to both of the two half axes of the ellipse. Referring to section A.3.2, this problem has four solutions, two corresponding to $P_1 = 1/2P_0(1 + sin[\Gamma(t)])$ and two corresponding to $P_2 = 1/2P_0(1 - sin[\Gamma(t)])$.

A.5.7 Modulator types

The electro-optic sensor is, from an optical point of view, an electro-optic amplitude modulator. Electro-optic amplitude modulators are generally grouped into two categories: transverse modulators where the electric modulation field and the direction of propagation of the optical wave are perpendicular; and longitudinal modulators where they coincide. In addition, a diagonal modulator will be described, for which the electric modulation field is applied at 45 degrees relative to the direction of propagation of the light wave. Figure A.7 shows an example of a transverse modulator and Figure A.8 shows a longitudinal modulator.

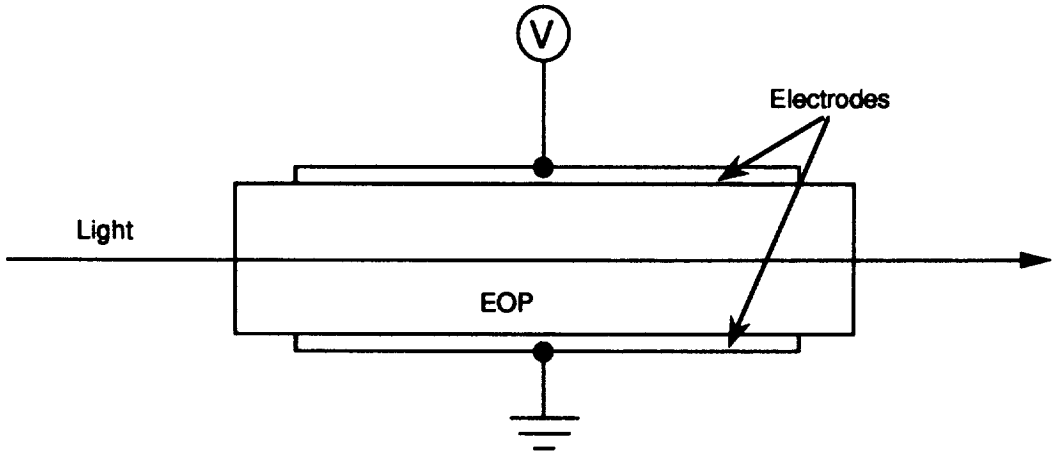


Figure A.7 Transverse modulator

E-field integrating sensors

The voltage V_{AB} between two points A and B is defined as the line integral

$$V_{AB} = -\int_A^B Edl \qquad (A.5.25)$$

where dl is the infinitesimal path length along an arbitrarily chosen path and **E** the electric field vector. A sensor which would be able to perform this line integral would provide an absolute measurement independent of local field inhomogeneities. The longitudinal and the diagonal electro-optic modulator would be able to perform this line integral, assuming that the electro-optic material is sandwiched between two electrodes, one at ground potential, and one at high potential, and the light beam is being allowed to traverse a path from one electrode to the other.

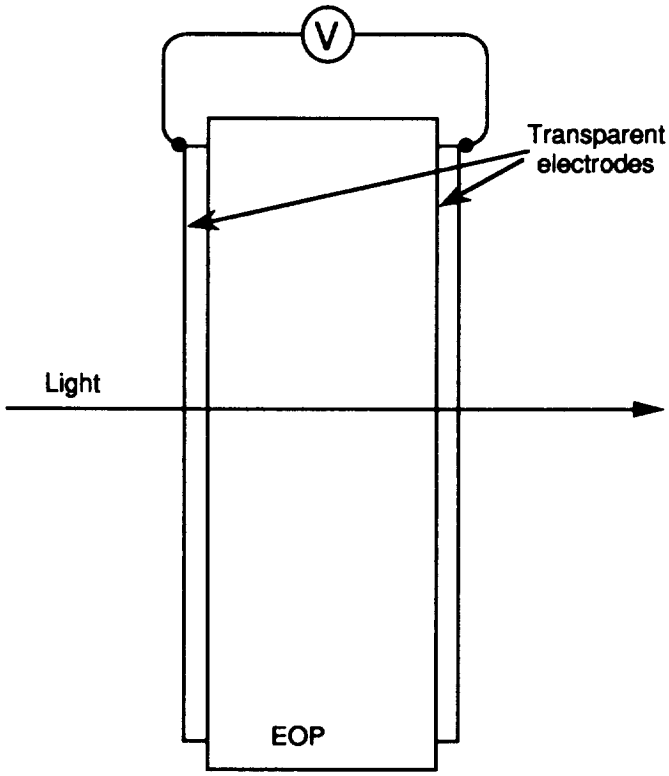


Figure A.8 Longitudinal modulator

For applications where the spacing between the two electrodes is not very large, the relative geometry of the electro-optic material and electrodes is fixed and the electro-optic material is in direct contact with the electrodes, the difference between the transverse and the longitudinal modulator is more academic than a matter of the accuracy of the device.

Case A: Direction of modulation field parallel to poling axis.

In this case, $E = (0, 0, E_z)$ and the index ellipsoid is represented by

$$(\frac{1}{n_0^2} + r_{13}E_z)x^2 + (\frac{1}{n_0^2} - r_{13}E_z)y^2 + (\frac{1}{n_0^2} + r_{33}E_z)z^2 = 1 \qquad (A.5.26)$$

The principal axes remain the same, and the indices of refraction are given by

$$\begin{aligned} n_x &= (\frac{1}{n_0^2} + r_{13}E_z)^{-\frac{1}{2}} \simeq n_0 - \frac{1}{2}n_0^3 r_{13}E_z \\ n_y &= (\frac{1}{n_0^2} + r_{13}E_z)^{-\frac{1}{2}} \simeq n_0 - \frac{1}{2}n_0^3 r_{13}E_z \\ n_z &= (\frac{1}{n_0^2} + r_{33}E_z)^{-\frac{1}{2}} \simeq n_0 - \frac{1}{2}n_0^3 r_{33}E_z \end{aligned} \qquad (A.5.27)$$

The electrically induced phase retardation will depend on the direction of propagation of the optical wave through the material.

Transverse modulation

Because of the rotational symmetry of the index ellipsoid around the z axis, the phase retardation in the transverse case will be independent of the orientation in the x and y directions. Consequently, if we calculate one solution, we will have them all.

The y axis is selected as the axis of propagation of the light wave. The analyzer is placed at an angle of 45 degrees relative to the z axis. The phase retardation is then given by

$$\Gamma = \frac{2\pi|n_x - n_z|L_y}{\lambda} \simeq \frac{2\pi n_0^3(r_{33} - r_{13})E_z L_y}{\lambda} \qquad (A.5.28)$$

where λ is the wavelength and L_y the dimension of the sample of material in the y direction.

When introducing the voltage V across the sample and the dimension L_z of the sample in the z direction, the phase retardation is given by

$$\Gamma \simeq \frac{\pi n_0^3(r_{33} - r_{13})V}{\lambda} \cdot \frac{L_y}{L_z} \qquad (A.5.29)$$

This is the configuration which has been used in most of the experiments with the electro-optic polymer, and the configuration used in the high voltage sensor. The implementation of this particular configuration is especially simple because the same electrodes can be used for application of the poling field and the modulation field.

Longitudinal modulation

Due to rotational symmetry around the z axis (the poling axis), a light wave propagating in this direction would not be exposed to any phase retardation. Consequently longitudinal modulation is not possible in this configuration.

Case B: Direction of modulation field perpendicular to poling axis

Because of the rotational symmetry around the z axis (poling axis), the orientation of the field in the x and y directions can be chosen arbitrarily. If a solution is calculated for one orientation, it will cover all other orientations.

The modulation field $E = (E_x, 0, 0)$ is chosen. The index ellipsoid is then represented by the equation

$$\frac{x^2 + y^2 + z^2}{n_0^2} + 2r_{13}zxE_x = 1 \qquad (A.5.30)$$

In order to bring the ellipsoid into the form of equation (A.5.24), a new set of principal axes x', y', and z' are chosen where the relation to the original principal system is

$$\begin{aligned} x &= x'\cos(45) - z'\sin(45) = \frac{x'}{\sqrt{2}} - \frac{z'}{\sqrt{2}} \\ y &= y' \\ z &= x'\sin(45) + z'\cos(45) = \frac{x'}{\sqrt{2}} + \frac{z'}{\sqrt{2}} \end{aligned} \qquad (A.5.31)$$

In the new principal system, the index ellipsoid is then represented by

$$(\frac{1}{n_0^2} + r_{13}E_x)x'^2 + \frac{y'^2}{n_0^2} + (\frac{1}{n_0^2} + r_{13}E_x)z'^2 = 1 \qquad (A.5.32)$$

This gives the indices of refraction given by

$$\begin{aligned} n_{x'} &= (\frac{1}{n_0^2} + r_{13}E_x)^{-\frac{1}{2}} \simeq n_0 - \frac{1}{2}n_0^3 r_{13}E_x \\ n_{y'} &= n_0 \\ n_{z'} &= (\frac{1}{n_0^2} + r_{13}E_x)^{-\frac{1}{2}} \simeq n_0 + \frac{1}{2}n_0^3 r_{13}E_x \end{aligned} \qquad (A.5.33)$$

Transverse modulation

The y axis is chosen as the axis of propagation of the optical light wave. The analyzer is placed with the direction of polarization parallel to the z axis. The induced phase retardation is then with the same terminology as in case A given by

$$\Gamma = \frac{2\pi|n_{z'} - n_{x'}|L_y}{\lambda} \simeq \frac{\pi n_0^3 r_{13}V}{\lambda} \cdot \frac{L_y}{L_x} \qquad (A.5.34)$$

where L_x is the dimension of the sample in the x direction.

Longitudinal modulation

Longitudinal modulation is not possible. This can be shown in the following manner. The index ellipsoid is studied in the plane $y' = 0$. The intersection of the ellipsoid and the plane is an ellipse represented by

$$(\frac{1}{n_0^2} + r_{13}E_x)x'^2 + (\frac{1}{n_0^2} + r_{13}E_x)z'^2 = 1 \qquad (A.5.35)$$

Longitudinal modulation would be possible only if the difference $n_z - n_y$ is a function of E_x. The index in the y direction is simply

$$n_y = n_{y'} = n_0 \qquad (A.5.36)$$

As a result, the problem is reduced to the determination of n_z. Due to the rotation transformation between the original principal system and the new principal system, solutions for n_z are to be found for values of x',z' where $x' = z'$. Substituting x', z' by a, the equation

$$(\frac{1}{n_0^2} + r_{13}E_x)a^2 + (\frac{1}{n_0^2} - r_{13}E_x)a^2 = 1 \qquad (A.5.37)$$

is solved, resulting in

$$a = \pm\frac{1}{\sqrt{2}} \cdot n_0 \qquad (A.5.38)$$

Making the same substitution in the coordinate transformation yields

$$n_z = \frac{|a|}{\sqrt{2}} + \frac{|a|}{\sqrt{2}} = n_0 \qquad (A.5.39)$$

Consequently longitudinal modulation is not possible because $n_z - n_y = 0$.

Case C: Modulation field diagonal to poling axis

By using x' or z' as the axis of propagation of the light wave, diagonal modulation can be obtained. It is diagonal because the electric modulation field is parallel to the x axis, which is oriented at an angle of 45 degrees to x' and z'.

By selecting z' as the axis of propagation, one obtains the electro-optically induced phase retardation

$$\Gamma = \frac{2\pi|n'_x - n'_y|L'_x}{\lambda} \simeq \frac{\pi n_0^3 r_{13} V}{\lambda} \cdot \frac{L'_x}{L_x} \qquad (A.5.40)$$

If x' is selected the result is the same except for the sign.

It was demonstrated that longitudinal modulation is not possible with either the modulation field perpendicular, or parallel to the poling axis. However, the case where the modulation field is applied at a skew angle relative to the poling axis has not been studied. As a result of symmetry around the poling axis, this problem can be reduced to the situation where E_z and one of the two other field components E_x,E_y are non-zero. If the modulation field $E = (E_x,0,E_z)$ is chosen, the equation for the index ellipsoid reduces to

$$(\frac{1}{n_0^2} + r_{13}E_z)x^2 + (\frac{1}{n_0^2} + r_{13}E_z)y^2 + (\frac{1}{n_0^2} + r_{33}E_z)z^2 + 2r_{13}zxE_x = 1 \qquad (A.5.41)$$

Due to symmetry considerations one would expect possible longitudinal solutions to be found, when the modulation field is applied at an angle 45 degrees relative to the poling axis. This problem has not been approached and remains to be studied.

A.6 Electro-optic Response

After the elaborate mathematical analysis it shall now be clarified which parameters can actually be measured by electro-optic means. When a sample of electro-optic material is tested the transmitted optical power is measured. When the hardware is optimized for maximum modulation, the transmitted optical power is given by one of the two equations

$$P_1 = \frac{1}{2}aP_0(1 - \sin[\Gamma(t)]) \tag{A.6.1}$$

$$P_2 = \frac{1}{2}aP_0(1 + \sin[\Gamma(t)]) \tag{A.6.2}$$

The sign depends on the adjustment of the 1/4 wave plate, the polarizer and the analyzer. Assuming the electrical modulation field is an AC field, the electro-optically induced retardation $\Gamma(t)$ can in both cases be determined by measuring the modulation index

$$m(t) = \frac{P_{ac}(t)}{P_{dc}} = \Gamma(t) \tag{A.6.3}$$

The electro-optic coefficients can be calculated from the equations in section A.5.7. using the relation $m = \Gamma$ where m is the RMS value of the modulation index, and Γ is the phase retardation calculated using the RMS value V of the voltage across the electro-optic polymer sample.

Example: The transverse modulator with poling axis and direction of modulation field parallel. This configuration, which has been used to test the samples of electro-optic polymer and is also used for the HV sensor, has the electro-optically induced phase retardation characterized by

$$\Gamma = \frac{\pi n_0^3(r_{33} - r_{13})V}{\lambda} \cdot \frac{L_y}{L_z} \tag{A.6.5}$$

Let l be the length of the sample in the propagation direction of the optical wave, and d be the width of the sample in the direction of the applied electric modulation field $E = V/d$. Substituting L_y by l and L_z by d in equation (A.6.5) yields the equation

$$\Gamma = \frac{\pi n_0^3(r_{33} - r_{13})V}{\lambda} \cdot \frac{l}{d} \tag{A.6.6}$$

When the modulation index $m = \Gamma$ is known, r_{33}, r_{13} can be calculated

$$r_{33} = \frac{3\Gamma\lambda d}{2\pi n_0^3 V l} \tag{A.6.7}$$

$$r_{13} = \frac{1}{3}r_{33} \tag{A.6.8}$$

Calculated numerical values for r_{33}, r_{13} are given in Section 2.2. Unfortunately, there is no easy way to verify the calculated values. The relative relationship between r_{33} and the material properties are expressed in (A.5.17)

$$r_{33} \sim N\mu\beta\frac{E_p}{5k_BT} \tag{A.5.17}$$

Unfortunately the absolute value is difficult to calculate from the material properties.